UNIVERSITAT DE BARCELONA

Conseqüències economètriques del grau de memòria en dades temporals

Ernest Pons Fanals

CONSEQÜÈNCIES ECONOMÈTRIQUES DEL GRAU DE MEMÒRIA EN DADES TEMPORALS

Ernest Pons Fanals

Tesi Doctoral dirigida pel Dr. Jordi Suriñach i Caralt en el marc del programa de doctorat "Economia i Territori. Mètodes Quantitatius" de la Universitat de Barcelona.

Departament d'Econometria, Estadística i Economia Espanyola.

Barcelona, Octubre de 1998.

A na Magda

Als meus pares i al meu germà

Voldria dedicar les primeres paraules d'aquesta Tesi Doctoral a aquelles persones que m'han ajudat a fer-la possible.

A en Jordi Suriñach li vull agrair tota la confiança i suport que m'ha manifestat durant aquests anys. Com a director de la Tesi, els seus comentaris, crítiques i suggeriments han estat decisius per la seva elaboració.

En Manuel Artís i en Jordi Suriñach, com directors del grup d'Anàlisi Quantitativa Regional, han sabut crear un ambient de treball i han posat els mitjans necessaris per l'elaboració d'aquesta Tesi Doctoral. També a la resta de companys del grup els voldria agrair el seu suport i paciència.

També ha estat fonamental per la realització d'aquesta Tesi l'existència d'un ambient de treball adequat per la recerca i la docència al Departament d'Econometria, Estadística i Economia Espanyola gràcies als seus directors, Manuel Artís primer i Miguel Angel Sierra després.

N'Enrique López també té part de responsabilitat en aquesta Tesi ja que ell, juntament amb en Jordi Suriñach, foren els que van despertar en mi l'interès per l'Econometria.

El camarada Tomás del Barrio ha tingut la paciència de discutir amb mi alguns dels resultats de la Tesi. Voldria des d'aquí fer palesa la meva més sincera gratitud. També he tingut la sort de comptar en molts aspectes de la vida acadèmica amb l'ajut i col·laboració d'en Miquel Clar amb qui estic també en deute.

Els professors Jordi Pons i Andreu Sansó van guiar les meves primeres passes en el món de la recerca durant els meus primers anys al Departament de manera que, sense ells, no hauria estat possible aquesta Tesi Doctoral.

Als meus amics i companys de despatx, Vicente Royuela, Pau Salsas i Carlos Soldevilla els vull agrair la seva paciència, l'haver aguantat les meves neures i angoixes i l'haver creat un ambient de treball i companyó immillorable. A més, el Vicente Royuela ha tingut la paciència de discutir hores i hores sobre algunes de les idees que es presenten a la Tesi.

Per altra banda, també voldria manifestar la meva gratitud al José Ramón García, la Rosina Moreno, la Lucia Quadrado i l'Esther Vayà que m'han suportat durant tantes hores.

També vull agrair a la Coloma Grandes i la Clara Mangrané la seva paciència i, sobretot, l'ajut que suposa treballar amb algú que mai, passi el que passi, et posarà mala cara .

Més enllà de l'àmbit acadèmic, vull manifestar el meu reconeixement als meus pares per ensenyar-me la diferència entre un home pobre i un pobre home i aquelles petites coses que són realment importants.

Finalment, vull posar de manifest el meu agraïment a na Magda que, amb la seva infinita paciència i el seu recolzament, ha fet possible aquesta Tesi.

A tots ells, dels que he rebut molt més del que mai els podré donar, dedico aquesta Tesi.

ÍNDEX

TERCERA PART: CONSEQÜÈNCIES ECONOMÈTRIQUES

INTRODUCCIÓ

La tesi doctoral que aquí es presenta no és fruit d'una idea aïllada sinó que n'és el resultat d'una evolució molt concreta en la recerca tant a nivell teòric com aplicat de l'autor al voltant de l'anàlisi i tractament de sèries econòmiques temporals en el sí del Grup d'Anàlisi Quantitativa Regional del Departament d'Econometria, Estadística i Economia Espanyola de la Universitat de Barcelona.

A nivell aplicat, l'experiència assolida en la modelització de sèries econòmiques, sobretot com a eina fonamental per a l'anàlisi de la conjuntura econòmica, demostra que les tècniques i models aplicats per obtenir prediccions a curt i mig termini de variables econòmiques poden influir en els resultats obtinguts de manera que, usant unes altres tècniques o models és possible arribar a resultats diferents. Així, l'elecció d'un o altre tipus de modelització, lluny de ser un tema marginal a la pràctica, pot condicionar la interpretació que hom acaba fent de la situació econòmica actual.

A nivell teòric la tesi és conseqüència d'un treball de recerca previ al voltant de l'aplicació de les tècniques estadístiques de l'anàlisi espectral a l'àmbit de l'Econometria. Aquestes tècniques, tot i la seva popularitat en l'anàlisi estadística, conformen un cos teòric poc usat en certs àmbits economètrics relacionats amb les sèries temporals fins no fa gaire. Durant els darrers anys, però, és habitual cada cop més recórrer al domini de les freqüències per analitzar certs aspectes de les sèries econòmiques.

Fruit d'aquestes dues fonts sorgeix la visió de que la utilització habitual dels tan populars models ARMA i ARIMA presenta algunes limitacions importants. Concretament, tot i que és innegable la gran utilitat d'aquests models, també és cert que hi ha algunes característiques d'aquests models que no acaben de correspondre's amb les característiques que presenten les observacions temporals d'algunes variables econòmiques.

Així, moltes variables econòmiques presenten una forta autocorrelació que sovint s'interpreta com la necessitat de prendre diferències de la variable mentre que al calcular la sèrie diferenciada s'observen símptomes de sobrediferenciació. Per altra banda, al tractar amb variables de freqüència superior a la anual, és molt habitual usar el model *línies aèries* pel càlcul de prediccions ja que sembla reproduir força bé les característiques de les variables econòmiques en el domini temporal, però si hom compara les característiques en el domini de les freqüències del model i de les dades s'observen diferències importants.

En definitiva, hom es troba sovint amb certes contradiccions entre la modelització suggerida per les característiques de la variable en el domini temporal i la suggerida per les característiques en el domini de les freqüències, mentre a nivell teòric ambdós tipus d'anàlisi no són més que les dues cares de la mateixa moneda. En aquest sentit, la motivació de la tesi que aquí es presenta és la d'avançar en la resposta a aquestes aparents contradiccions. Així l'objectiu principal de la tesi és mostrar com algunes d'aquestes aparents contradiccions desapareixen quan es té en compte el que considerem una característica fonamental de les sèries temporals i que hem anomenat a la tesi *memòria*[1].

A nivell intuïtiu es pot interpretar aquest concepte com la màxima distància temporal a través de la qual es manté una certa dependència entre els valors de la variable. De fet, la idea de *memòria* és present de manera indirecta a l'anàlisi estadística de sèries temporals d'ençà el desenvolupament del concepte d'estacionarietat en forma de restriccions per a que els moments mostrals siguin bons estimadors dels moments poblacionals. En aquest sentit, posar restriccions a la *memòria* d'una sèrie temporal consisteix a limitar el nivell d'informació que les observacions passades aporten a la predicció de les observacions futures i la manera més popular de formalitzar aquestes restriccions és a través del concepte d'*ergodicitat*.

Però no sols a l'anàlisi teòrica de les sèries temporals s'hi pot trobar la idea de *memòria* sinó que també a l'anàlisi econòmica hi és present tot i que en aquest cas és més habitual usar el mot *persistència*. En aquest àmbit l'existència d'una major o menor *persistència* d'una variable fa referència a la major o menor influència temporal en l'evolució futura de la variable davant l'aparició d'un shock en el valor present de la variable.

---

[1] S'ha optat per mantenir el mot en cursiva mentre no se'n doni una definició formal.

Com exemple més típic d'aquest interès és de destacar l'àmplia discussió generada al voltant del grau de *persistència* que presenta l'evolució del Producte Interior Brut de diferents països, per raons de política econòmica.

De totes maneres, tot i aquests antecedents cal destacar que a la tesi, a diferència dels plantejaments anteriors, es presenta un concepte de *memòria* molt més general que permet classificar les sèries temporals com a pertanyents a vàries categories segons quin sigui el seu grau de *memòria.* Sota aquest plantejament considerem la *memòria* com una qualitat que conjuntament amb l'estacionarietat o no estacionarietat i la possible presència d'estacionalitat determina el tipus de tractament i tècniques que cal utilitzar.

De manera habitual, la valoració dels models paramètrics o semiparamètrics usats per la representació de dades econòmiques temporals es fa contrastant si es pot acceptar que els residus són soroll blanc. Aquest plantejament es basa, de manera implícita, en que l'absència de *memòria* en una sèrie temporal és equivalent a que sigui soroll blanc. En tot cas creiem que és important anar mes enllà en aquesta idea i considerar el grau de *memòria* com una característica no paramètrica fonamental en les dades temporals. Així, per validar un model per les dades cal comprovar que la *memòria* de les dades és la mateixa que la *memòria* del model.

Sovint, en el tractament economètric de variables econòmiques a partir d'informació de caire temporal no es té en compte el tipus de *memòria* que contenen aquestes dades o, si es té en compte s'imposen restriccions que poden ser excessives. Per tant, la tesi neix de la preocupació pels errors que pot provocar no considerar de manera adequada el tipus de *memòria* i especificar models que no reprodueixen de manera adequada les característiques temporals d'aquestes variables.

Així, tot i que aquests problemes d'especificació i selecció de models per sèries temporals són part de l'estadística matemàtica, cal remarcar que l'interès final de la tesi es troba en les conseqüències economètriques del concepte de *memòria* tot i que per assolir aquest objectiu, calgui tractar en alguns moments al llarg de la tesi aspectes purament estadístics.

Per tant, l'objectiu final de la tesi és el de valorar la rellevància del concepte de *memòria* de cara a l'anàlisi economètrica de dades temporals. Aquest objectiu inclou dos aspectes igualment importants, d'una banda és necessari detectar quines poden ser les conseqüències negatives que pot tenir la utilització de les eines economètriques més habituals sense tenir en compte el tipus de *memòria* de les dades. I d'altra banda, un cop detectades aquestes conseqüències, cal avançar en la recerca d'instruments adequats per evitar aquests problemes.

Òbviament, aquest objectiu supera en molt el que és possible abordar, per raons d'espai, en una tesi doctoral. Per aquest motiu és fonamental emmarcar l'anàlisi d'aquestes conseqüències en un context teòric adequat per situar tant els resultats assolits a la tesi com altres resultats previs i resultats que es puguin derivar d'una recerca posterior.

Per això, s'ha posat especial èmfasi a la tesi a donar un conjunt de definicions, el més generals possibles, sobre el concepte de *memòria* que *a)* siguin especialment adequades per analitzar les implicacions d'aquest concepte des d'un punt de vista economètric, *b)* permetin relacionar els resultats de la tesi amb resultats previs obtinguts en altres treball i *c)* siguin el marc adequat per continuar amb aquesta línia de recerca. L'establiment d'aquest marc de referència creiem que és, al marge dels resultats més concrets obtinguts, una de les aportacions importants de la tesi.

En aquest sentit, creiem important destacar que el plantejament proposat a la tesi obra una àmplia via de recerca futura que creiem pot portar a resultats importants per a l'anàlisi economètrica. És per aquest motiu que algunes de les aportacions que es presenten a la tesi s'han de considerar, més que com a resultats definitius, aportacions que incorporen alguns elements novedosos i han de permetre obtenir en el futur nous resultats en l'anàlisi economètrica de sèries temporals.

Per raons d'espai, a la tesi es parteix de la base d'un coneixement bàsic de la majoria de conceptes i tècniques de l'anàlisi de sèries temporals com són la modelització ARMA, anàlisi temporal i espectral, teoria de la integració, tècniques d'extracció de senyals, teoria de la informació, teoria de la predicció i altres tòpics de l'anàlisi estadística de sèries temporals. En aquest sentit, la tesi no conté de manera estricta les definicions formals de tots els conceptes usats sinó que sols es presenten les definicions novedoses i aquelles que, tot i ser presents a la literatura, se'n poden trobar varies definicions. En els casos en que els conceptes són prou estàndards i estan prou consolidats no s'entra en la seva discussió.

Per assolir els objectius plantejats s'ha dividit la tesi en tres grans blocs. En primer lloc, és necessari caracteritzar formalment el concepte de *memòria*. Per tant, la primera part de la tesi es centra en la presentació del concepte posant especial èmfasi en la seva relació amb altres treballs i línies de recerca. Així, com a primer objectiu es presenten diferents conceptes de *memòria* desenvolupats al llarg del temps per diferents autors, així com diferents mètodes proposats a la literatura per mesurar el que aquests autors entenen per *memòria*.

L'objectiu del capítol 1 és introduir de manera simple, tot i que poc formal, el concepte de *memòria* i mostrar la seva rellevància per a les aplicacions economètriques. A més, es posa especial interès a realitzar una breu referència històrica dels diferents treballs relacionats amb aquest concepte a nivell economètric.

Aquesta revisió és necessària ja que, com s'ha comentat, es tracta d'un concepte que es troba de manera molt dispersa i sota enfocaments bastant diferents a la literatura, tant estadística com economètrica. En particular, es posa especial èmfasi en les aplicacions economètriques del concepte de *memòria llarga* que s'han desenvolupat en l'última dècada.

Tot i que considerem el plantejament inclòs en capítols posteriors de la tesi nou per la seva sistematicitat i generalitat, el concepte de *memòria* d'una sèrie temporal ha aparegut de manera freqüent i bastant recurrent a l'anàlisi de sèries temporals en diferents entorns. Així, al capítol 2 es presenten un conjunt d'instruments que anomenem *clàssics* usats a la literatura economètrica per mesurar el grau de *persistència* de variables econòmiques. La característica d'aquestes mesures és el fet que es tracta de mesures prèvies a la definició formal dels diferents tipus de *memòria* però que permeten un seguit d'aproximacions molt interessants sobre aquest concepte.

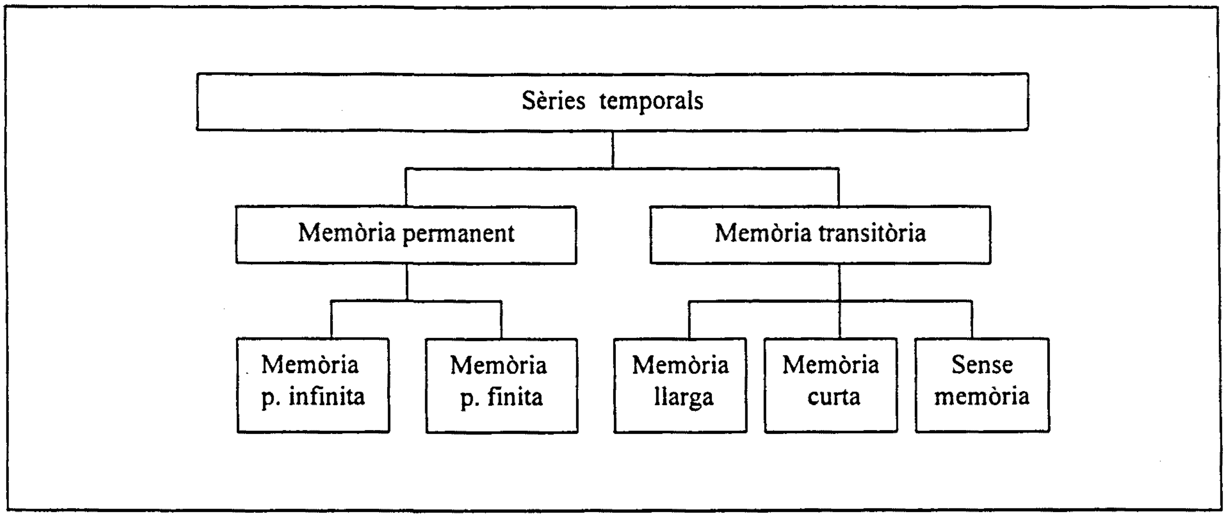
Un cop presentat el concepte de *memòria*, la seva interpretació i alguns instruments per la seva mesura, cal donar una definició formal i precisa dels diferents tipus de *memòria* que pot presentar una sèrie temporal. Per fer-ho hi ha diferents possibilitats que, tot i molt similars, no són completament equivalents. Així, es dedica el capítol 3 a la presentació, comparació i discussió d'aquestes possibles definicions.

De totes maneres, aquestes definicions no són ni equivalents entre sí, ni adequades algunes d'elles per la seva utilització en un àmbit economètric. Per això, ha calgut seleccionar la classificació més adequada pels objectius de la tesi. Per tant, atès que es tracta d'una elecció que pot ser sotmesa a discussió es posa especial interès en la justificació d'aquesta elecció.

Més concretament es presenta una classificació en forma d'arbre format per diferents nivells. En primer lloc i en el nivell superior, es poden dividir les sèries temporals entre sèries amb *memòria permanent* i sèries amb *memòria transitòria*. En segon lloc, les dades amb *memòria permanent* es divideixen entre sèries amb *memòria permanent infinita* o sèries amb *memòria permanent finita*. Pel que fa a les sèries amb *memòria* transitòria es poden dividir entre sèries sense *memòria*, sèries amb *memòria curta* i sèries amb *memòria llarga*.

La divisió entre *memòria permanent* i *memòria transitòria* està associada, a grans trets i sense ser-ne equivalent, a la divisió entre estacionarietat i no estacionarietat mentre que la divisió entre *memòria infinita* i *memòria finita* està associada al concepte de reversió cap a la mitjana. Pel que fa a la distinció entre *memòria curta* i *memòria llarga*, és pràctic interpretar la primera com aquella que presenten els modes ARMA mentre que la segona correspon a una situació estacionària amb una major *memòria*. Com a conseqüència d'aquesta classificació, s'observa que les tècniques economètriques habituals estan basades en el supòsit que les dades són de *memòria permanent infinita*, de *memòria curta* o *sense memòria* sense considerar les altres dues possibilitats.

Quadre 0.1. Classificació de les sèries temporals segons el tipus de *memòria.*



Un cop definits els diferents tipus de *memòria*, abans d'abordar l'anàlisi de les conseqüències economètriques d'aquesta diversitat és necessari desenvolupar una base estadística que permeti avançar en aquesta anàlisi. Per això s'inclou a la tesi una segona part que es centra en els aspectes més de tipus estadístic relacionats amb el concepte de *memòria.* Així, es presenten un conjunt de models estadístics i un conjunt de tècniques que permeten la realització d'inferència estadística al voltant d'aquests models.

La disponibilitat d'aquests models és necessària per abordar després la valoració de les conseqüències economètriques ja que és força difícil fer-ho en un context massa genèric. Per això, si hom disposa d'un conjunt de models que permetin reproduir els diferents tipus de *memòria*, és possible analitzar com afecta a l'anàlisi economètrica la utilització de variables generades per cada un d'aquests models.

És per aquest motiu que es dedica el capítol 4 a la presentació i comparació de diferents possibilitats, algunes ja proposades per altres autors i d'altres més noves que hi ha per a la modelització de sèries temporals segons el tipus de *memòria* que presenten. Donat que la modelització de sèries temporals que presenten *memòria curta*, en termes de les definicions del capítol 3, és un problema sobre el que ja es disposa de gran quantitat de resultats, es dedica aquest capítol de manera específica a diferents possibilitats per la modelització de sèries temporals que contenen *memòria llarga.*

Pel que fa a les sèries amb *memòria permanent*, com són no estacionàries cal usar alguna transformació prèvia per transformar-les en estacionàries amb *memòria curta* o bé *memòria llarga* i usar models adequats a cada un d'aquests casos.

Un cop s'ha definit al capítol 3 el concepte de *memòria* i s'han presentat al capítol 4 un seguit de models estadístics per representar sèries amb *memòria llarga,* cal disposar d'algun mètode pràctic d'identificació que permeti esbrinar, a partir d'un conjunt d'observacions, el tipus de *memòria* que conté una certa variable.

Això és important ja que algunes de les conclusions de la tesi fan referència a que el tipus d'eines adequades pel tractament d'observacions temporals depèn del tipus de *memòria* d'aquestes dades. D'aquí que sigui tan important disposar de tècniques que permetin la identificació pràctica del tipus de *memòria* problema al que es dedica el capítol 5 de la tesi.

La disponibilitat d'aquests models porta directament associat el problema de la seva estimació. Òbviament, si aquests models s'han d'usar com a eines economètriques, és necessari conèixer quina és la millor manera de realitzar inferència al voltant d'aquests models.

A més, en molts d'aquests models el tipus de *memòria* està associat amb un paràmetre concret del model, de manera que la realització d'inferència sobre aquest paràmetre és un mètode per la identificació pràctica del tipus de *memòria*, alternatiu als presentats al capítol 5.

Com el problema de l'estimació de models per variables amb *memòria curta* és un problema que es troba bàsicament resolt i pel que ja es disposa de força resultats, sols té interès analitzar el problema de l'estimació de models per variables amb *memòria llarga.* Per aquest motiu els capítols 6 i 7 estan centrats en el problema de l'estimació de models de *memòria llarga* per dades econòmiques.

A grans trets, atès que es tracta d'un concepte no paramètric, hi ha dos vies per a la seva estimació, seguint una aproximació paramètrica o una aproximació semiparamètrica. Així, es dedica a la primera opció el capítol 7, mentre que el capítol 6 es centra en l'estimació paramètrica dels models presentats al capítol 4. Com ja s'ha senyalat, l'estimació d'aquests models no deixa de ser més que una via paramètrica de mesurar la *memòria* d'una sèrie temporal alternativa a les mesures presentades al capítol 2. La diferència es troba en que les anteriors no es basen en una definició precisa del concepte *memòria.*

Aquesta dicotomia entre mesures no paramètriques i models paramètrics està molt relacionada amb un aspecte pràctic molt important: l'elecció del tipus de modelització més adequada per a una sèrie temporal concreta. D'una banda, hi ha l'opció d'usar diferents models i valorar després els resultats que s'obtenen, mentre que l'altra possibilitat és usar *a priori* algun instrument no paramètric que permeti classificar les característiques de les dades analitzades i escollir el tipus de modelització més adient. De fet, per motius que s'analitzen al llarg de la tesi, la discriminació entre els diferents tipus de *memòria* no és un problema senzill.

Finalment, i assolint l'objectiu principal, la tercera part de la tesi es centra en l'anàlisi de les conseqüències economètriques que pot tenir la no consideració de que les dades temporals poden presentar diferents graus de *memòria* i, per tant, la importància de tenir en compte aquest concepte. Donat que entre aquestes conseqüències n'hi ha moltes que han estat analitzades de manera més o menys indirecta a la literatura economètrica, l'anàlisi es centra en aquelles conseqüències menys analitzades.

Per una banda, el capítol 8 es centra en les conseqüències del tipus de *memòria* a nivell d'anàlisi univariant. Algunes de les conseqüències més importants que poden tenir els errors en l'especificació del tipus de *memòria* són les següents:

- Es produeixen problemes greus de biaix en els estadístics més habituals tant en el domini temporal com en el domini de les freqüències. Al llarg de la tesi s'analitzen aquests problemes tant en el domini temporal com en el domini freqüencial, i es presenten algunes propostes per tal de minimitzar aquests problemes.

- S'analitzen també els problemes associats amb la selecció i identificació de models paramètrics, per exemple models ARMA, quan es produeix un error en el diagnòstic del tipus de *memòria* de les dades. En particular, s'analitzen els problemes que presenten els tests d'arrels unitàries més habituals quan les dades presenten *memòria permanent finita.*

- L'especificació errònia de models paramètrics per dades temporals també provoca importants problemes quan l'objectiu d'aquesta modelització és el càlcul de prediccions. En aquest sentit, s'avança en l'anàlisi de mètodes adequats pel càlcul de prediccions segons el tipus de *memòria* de les dades.

- Finalment, la modelització estocàstica de sèries temporals ha adquirit recentment una importància addicional ja que els mètodes més moderns per a l'estimació de components no observables s'han desenvolupat sota el supòsit que aquestes components es poden representar a través de models ARIMA. Segons quin sigui el tipus de *memòria* de les dades, aquesta hipòtesi pot tenir greus conseqüències pel que fa a les propietats de les components estimades.

Per altra banda, el capítol 9 es centra en les conseqüències que té no tenir en compte el tipus de *memòria* per l'estimació de models de regressió. El més important d'aquests problemes és la possibilitat d'obtenir regressions espúries segons el tipus de *memòria* de les dades. Aquest fenomen ja conegut pel cas de dades no estacionàries, en particular en sèries integrades, no és exclusiu d'aquest tipus de *memòria.*

A la tesi es demostra que a partir de dades que presenten *memòria permanent finita* i, fins i tot, amb dades estacionàries que presenten *memòria llarga* també és molt probable obtenir regressions espúries. Com a solució a aquest problema es proposa utilitzar un mètode d'estimació basat en el domini de les freqüències que minimitza el perill d'obtenir regressions espúries.

Simultàniament a l'anàlisi de cada una d'aquestes conseqüències, i com a objectiu addicional de la tesi, es proposa una metodologia i un conjunt d'instruments que permeten pal·liar en la mesura del possible alguns d'aquests problemes. Òbviament, molts d'aquests problemes resten pendents d'una recerca posterior més amplia. A més, alguns dels instruments proposats obren línies de recerca molt interessants per a treballs posteriors.

Per acabar, al capítol 10 es realitza un resum dels resultats més importants presentats i es presentes les conclusions més importants que d'aquests se'n deriven. Així mateix, s'analitzen un conjunt d'aspectes relacionats amb la *memòria* en dades temporals que queden pendents d'una recerca posterior més amplia.

PRIMERA PART

CONCEPTE

Capítol 1

LA NOCIÓ DE *MEMÒRIA* EN SÈRIES TEMPORALS

1.1. Introducció

Moltes de les dades usades a l'anàlisi economètrica estan organitzades en forma de sèrie temporal. En aquestes circumstàncies no és raonable mantenir el supòsit que aquestes dades són realitzacions de variables aleatòries independents i idènticament distribuïdes. És doncs per això que l'anàlisi de dades temporals ha generat tant interès en els econòmetres. Per aquest motiu s'han desenvolupat tot un conjunt de models paramètrics, semiparamètrics i no paramètrics amb la intenció de descriure algunes característiques de les sèries temporals, especialment l'estructura de dependència que presenten aquest tipus de dades.

Aquests models s'han usat per fer prediccions, per mesurar la dependència entre variables econòmiques, per contrastar hipòtesis econòmiques, per mesurar la persistència, per descriure variables latents o no observables i moltes altres aplicacions, però des d'un punt de vista economètric, un dels aspectes fonamentals de l'anàlisi estadística de les sèries temporals consisteix en que permet derivar teoremes sobre límits centrals i sobre la consistència dels diferents estimadors usats habitualment en l'estimació de models economètrics, malgrat la dependència que presenten habitualment les dades temporals.

Així, l'anàlisi recent de les sèries temporals des d'una vessant economètrica s'ha centrat en relaxar el supòsit d'independència entre les observacions, mantenint el d'igual distribució[1]. En aquest context, la funció d'autocorrelació s'ha mostrat com una eina fonamental per al tractament de dades temporals. Aquesta funció es pot estimar de manera no paramètrica tant en el domini temporal com freqüencial, i és habitual la seva utilització en sèries llargues per descriure l'estructura de dependència de les dades. En canvi, quan es disposa d'un nombre no gaire gran d'observacions, fet molt comú quan es tracta de variables econòmiques, és més habitual la utilització de models paramètrics, sobretot models ARMA. Si al supòsit

[1] Que en aquest cas es coneix amb el nom d'estacionarietat estricta.

d'estacionarietat[2] s'hi afegeix el d'ergodicitat, això permet garantir que els moments mostrals són estimadors consistents dels moments poblacionals.

De fet, el tipus de dependència temporal juga un paper fonamental en les propietats dels diferents estadístics. Donada l'especial simplicitat del tractament dels models ARMA, no és d'estranyar que s'hagin desenvolupat tot un seguit de resultats economètrics sota la hipòtesi de que els termes de pertorbació d'un model economètric són susceptibles d'una representació d'aquest tipus.

Així, en el context dels models ARMA, quan es suposa que les innovacions del procés són independents i igualment distribuïdes (*iid*) o bé es suposa que són una diferència de martingales (*dm*), és possible demostrar que els estimadors habituals, tant els paramètrics com els no paramètrics són a més, asimptòticament normals. De manera encara més general, tot procés lineal amb innovacions *iid* o *md* compleix un teorema central del límit que estableix que el rati de convergència és el mateix que en el cas en que les dades són *iid* sempre que els coeficients de la representació en forma de mitjana mòbil decreixin prou ràpid.

La majoria de models economètrics tenen la característica bàsica que hi ha un conjunt de variables, que es consideren rellevants, que són ortogonals o independents d'un altre grup de variables no observables o pertorbacions, i que depenen a més d'aquestes pertorbacions d'un conjunt de variables observables i d'un conjunt de paràmetres. En aquest context, els resultats de la inferència clàssica en models de regressió lineal amb pertorbacions *iid* continuen sent vàlids si els regressors són estocàstics sempre que siguin independents dels termes de pertorbació. Així mateix, tampoc hi afecta el fet de que aquests regressors presentin una estructura temporal sempre que es mantingui la independència amb el terme de pertorbació.

Per models més generals com sistemes d'equacions simultànies o models de retards distribuïts molts dels resultats asimptòtics de consistència són encara vàlids segons l'estructura temporal

[2] Al llarg de la tesi, si no s'especifica el contrari, s'utilitza el concepte estacionarietat per referir-se a l'estacionarietat de segon ordre o estacionarietat dèbil.

de les variables implicades. En particular, quan l'estructura de dependència és la d'un model ARMA es garanteixen moltes de les propietats dels estimadors. En tot cas, la variància asimptòtica dels estimadors sí que es veu sovint afectada per la presència d'autocorrelació en les pertorbacions i en aquests casos s'han proposat varietat d'estimadors alternatius asimptòticament eficients (per exemple Cochrane i Orcutt, 1949 o Hannan, 1963).

De fet, la dependència de molts dels resultats teòrics a la teoria economètrica respecte del supòsit de que les dades es poden representar a través d'un model ARMA destaca la rellevància tant teòrica com aplicada de la idoneïtat d'aquest supòsit. En aquest sentit, els models ARMA han estat molt utilitzats per modelitzar sèries econòmiques sota la justificació del Teorema de representació de Wold quan s'accepta que aquestes són estacionàries. Recordi's que aquest teorema demostra que la part estrictament no determinista de tot procés estocàstic estacionari es pot representar mitjançant una mitjana mòbil infinita.

Així, a partir de l'existència d'una representació com mitjana mòbil infinita s'aprofita la possibilitat de trobar una aproximació prou bona a aquest model infinit a través d'un model ARMA(p,q) si s'utilitzen ordres p i q prou grans. Val la pena destacar que l'estratègia de modelització més popular, la proposada a l'obra de Box i Jenkins (1970) va més enllà d'aquests resultats i es basa en el supòsit que n'hi ha prou amb usar ordres p i q petits[3].

Però els models ARMA sols estan justificats sota el supòsit d'estacionarietat i moltes de les dades econòmiques presenten forta evidència de no estacionarietat. Com a mínim, és habitual a l'anàlisi de variables econòmiques trobar correlogrames amb coeficients molt important fins i tot per retards grans. Usant una altra terminologia, és habitual que moltes variables econòmiques presentin una *persistència* molt important.

Aquest fenomen no és coherent amb les característiques dels models ARMA i per aquest motiu s'han intentat usar models inestables per descriure aquesta característica. El més senzill és la

[3] Condició que es coneix amb el nom de principi de parsimònia.

utilització d'un AR(1) amb paràmetre superior o igual a la unitat en valor absolut. En el cas explosiu, en que el paràmetre autoregressiu és estrictament superior a la unitat, White (1958) i Anderson (1959) han mostrat que la distribució límit de l'estimador MQO del paràmetre autoregressiu pot seguir tant una distribució Normal com Cauchy. Chan i Wei (1988) estenen aquesta teoria a un conjunt més ampli de models.

De fet, els econòmetres s'han centrat bàsicament en el cas d'arrels unitàries, popularitzades en el treball de Box i Jenkins (1970) que posa de manifest la utilitat de considerar models en que les dades un cop diferenciades es comporten segons un model ARMA. White (1958) mostra que en aquest cas la distribució límit de l'estimador del paràmetre autoregressiu és una funció no lineal i no estàndard d'un moviment brownià. Desenvolupaments posteriors d'aquesta teoria es poden trobar a Rao (1978), Dickey i Fuller (1979). La teoria s'ha ampliat després per considerar una major varietat de models, el comportament estacional, els canvis estructurals i altres tòpics de la recerca economètrica actual. La teoria de distribucions límit és notable en aquest context per la seva diversitat. En alguns casos aquesta distribució és normal, però en la majoria depèn de moviments brownians.

La motivació principal per l'estudi d'aquestes distribucions ès el desenvolupament de tests d'arrels unitàries. Com no es poden trobar valors crítics en les taules estadístiques estàndards, trobar aquests valors és una tasca intensiva de càlcul amb ordinadors. Una solució front a aquesta varietat de situacions és la utilització de superfícies de resposta. Les regions crítiques també varien al usar altres test com el de Wald, la raó de versemblança o els multiplicadors de Lagrange.

En tot cas, tot i que la noció de forta *persistència* de les dades s'associa habitualment amb el concepte d'integrabilitat es mostra al llarg de la tesi que sols n'és un cas molt particular. Dit d'un altre manera, no és correcte restringir l'anàlisi de sèries econòmiques al supòsit de que aquestes sèries, un cop eliminats els possibles elements deterministes, són o bé models ARMA o bé models ARIMA perquè aquest conjunt de models no recullen totes les possibilitats.

Concretament, tots els models ARMA(p,q) amb ordres p i q no massa grans, presenten funcions d'autocorrelació simple i parcial que decreixen molt ràpidament. Així, s'acostuma a afirmar que els models ARMA tenen *memòria curta* perquè un *shock* en un determinat moment del temps té efectes posteriors molt limitats en el temps. Això és equivalent, en termes de predicció, a que la predicció dinàmica del comportament futur de la sèrie ràpidament convergeixi a la seva esperança. Però, al menys teòricament, un procés pot ser estacionari però presentar unes autocorrelacions que decreixin molt lentament cap a zero.

Més concretament, mentre que per a que una sèrie sigui estacionària és suficient que aquests coeficients d'autocorrelació siguin quadràticament sumables, en el cas dels models ARMA es compleix a més que aquests coeficients són absolutament sumables[4].

De fet, hi ha forta evidència empírica de sèries temporals en diferents ciències físiques[5] que presenten unes autocorrelacions que decreixen molt lentament, és a dir, que mostren una forta dependència entre observacions molt allunyades en el temps sense que sembli raonable rebutjar la hipòtesi d'estacionarietat. D'aquesta observació sorgeix el concepte de *memòria llarga* per distingir aquesta possibilitat, per exemple, de la que caracteritza als models ARMA.

Així, i de manera informal, es pot situar el concepte de *memòria llarga* com un nivell intermedi entre les característiques dels models ARMA i les dels models ARIMA. De fet, en

---

[4] Formalment, si ρ_k són els coeficients d'autocorrelació d'una sèrie estacionària X_t, es pot demostrar a partir de la condició d'estacionarietat de segon ordre que $\sum \rho_k^2 < \infty$ i es diu que la successió de coeficients és quadràticament sumable. Si la sèrie es pot representar com un model ARMA, es pot demostrar que $\sum \left|\rho_k\right| < \infty$ i es diu que la successió és absolutament sumable. La condició de sumabilitat absoluta és una propietat més forta ja que tota successió absolutament sumable també és quadràticament sumable mentre que la implicació contraria no és certa.

[5] La disponibilitat en aquests disciplines científiques de sèries temporals formades per un gran nombre d'observacions permet comprovar de manera fiable la presència d'aquesta correlació entre obervacions molt allunyades en el temps. En canvi, les limitacions pel que fa a la longitud de la majoria de sèries de l'àmbit econòmic fa difícil la comprovació empírica d'aquesta característica.

una sèrie que conté alguna arrel unitària té sentit, almenys de manera intuïtiva, referir-se a *memòria infinita* ja que la correlació entre dades cada cop més allunyades en el temps no tendeix cap a zero, és a dir, la dependència no desapareix amb el pas del temps.

En tot cas, és important donar una definició més precisa d'aquests conceptes i aquesta és justament una de les tasques abordades al llarg de la tesi, concretament al capítol 3. Però deixant de banda, de moment, aspectes més formals, s'ha observat com moltes de les sèries temporals econòmiques presenten característiques molt similars a les que s'entén que presentarien sèries amb *memòria llarga*, de manera que és important considerar les conseqüències que pot tenir la presència d'aquestes característiques per a la inferència economètrica. Especialment importants són les conseqüències d'aplicar tests d'arrels unitàries i modelitzar a continuació, a través de models ARMA o ARIMA, dades que presenten *memòria llarga* o *memòria permanent finita* ja que s'està cometent un error d'especificació del model. No tant per la modelització en sí, sinó pel que fa a les conseqüències posteriors a l'hora d'estimar relacions economètriques entre varies d'aquestes variables.

1.2. Breu revisió històrica

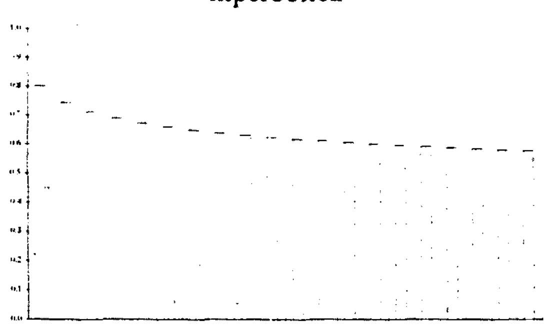
De fet, l'evidencia empírica de la presència d'una forta autocorrelació en sèries de diferents àmbits científics és molt antiga, molt anterior als més populars models ARMA o ARIMA. Però la modelització i la formalització estadística d'aquest concepte és molt més recent. Tant, que encara hi ha molts aspectes que resten per estudiar i, de fet, en aquest moment s'hi produeix una àmplia recerca.

Newcomb (1886) i Jeffreys (1939) troben evidència d'aquesta forta correlació en dades astronòmiques. Així també Student (1927) amb dades químiques. Smith (1938) i Wittle (1956) detecten també aquest fenomen en gràfics de dades agrícoles.

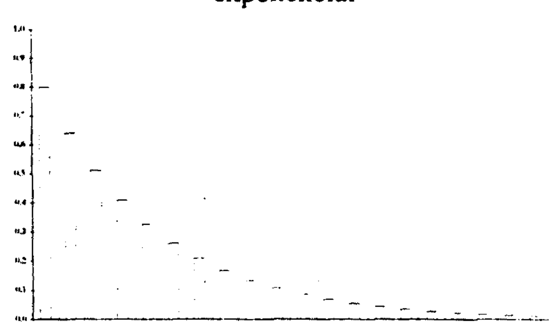
Però l'aplicació pionera més coneguda amb diferència és a la hidrologia. Concretament Hurst (1951) ho detecta en el comportament de les onades en un dipòsit i després a sèries temporals de molts altres àmbits (Hurst, 1956). A més, proposa un estadístic, el *rang reescalat*[6] per l'anàlisi d'aquest fenomen. Les propietats d'aquest estadístic han estat àmpliament estudiades per Mandelbrot i Wallis (1968 i 1969b), Mandelbrot (1972 i 1975a) i McLeod i Hipel (1978). A Lawrance i Kottegoda (1977) i Kashyap i Lapsa (1984), es destaca la importància de la *memòria llarga* en diferents àmbits de la recerca científica i es recopilen varis models usats per modelitzar aquest tipus de comportament.

En aquests models és habitual el supòsit que les autocovariàncies presenten un decreixement de tipus hiperbòlic i és molt habitual suposar que asimptòticament les autocorrelacions es poden representar segons $|\gamma_k| \sim k^{2\delta-1}$ amb $0 \le \delta < 1/2$. Recordi's que en models ARMA el decreixement de les autocorrelacions segueix un esquema exponencial que es pot representar com $|\gamma_k| \sim C^k$. Als gràfics 1.1 i 1.2. es mostren exemples de cada tipus de decreixement i es pot comprovar la gran diferència que hi ha entre els dos, tot i tenir un primer coeficient d'autocorrelació comú.

Gràfic 1.1. Correlograma amb forma hiperbòlica



Gràfic 1.2. Correlograma amb forma exponencial



[6] Conegut també com estadític R/S i analitzat a l'apartat 2.5 de la tesi.

De fet, ja molts anys abans, Kolmogorov (1940) havia introduït una família de processos estocàstics a temps continu, que presenten autocovariàncies amb estructura hiperbòlica i que s'han usat bastant a la física teòrica[7]. A partir d'aquesta estructura d'autocovariàncies hiperbòlica, i sota el supòsit de normalitat, Mandelbrot i Van Ness (1968) desenvolupen el procés estocàstic conegut com soroll gaussià fraccional, presentat a l'apartat 4.3. de la tesi.

Pel que fa més específicament a la ciència econòmica, a la dècada del anys seixanta apareix força interès per l'estimació no paramètrica de l'espectre de les sèries econòmiques. En moltes sèries l'espectre presenta un salt al voltat de la freqüència zero i decreix ràpidament a la resta de les freqüències[8]. Aquest fenomen fou identificat per Granger (1966) com *"the typical spectral shape of an economic variable"*. Aquest aspecte és consistent amb la presència d'una o més arrels unitàries però moltes d'aquestes sèries un cop diferenciades semblen presentar un espectre nul a la freqüència zero, fet incompatible amb la presència d'arrels unitàries.

En canvi, aquest fet és compatible amb una estructura d'autocorrelacions hiperbòlica, per exemple la dels models proposats per Kolmogorov i Mandelbrot. A Mandelbrot (1969, 1972 i 1973) es justifica la importància d'aquest tipus de models en economia i es defensa la utilització de l'estadístic R/S. De fet, el mateix Mandelbrot és molt crític amb les tècniques habituals en aquella època en l'anàlisi de dades temporals econòmiques. Concretament, proposa interpretar la densitat espectral estimada amb molta cura quan les sèries econòmiques mostren gran *persistència*, fet que considera habitual.

Més tard, la utilització de l'estadístic R/S és depurada i formalitzada des d'un punt de vista estadístic per Lo (1991). A més, són també dignes de destacar els exercicis de Montecarlo realitzats pel mateix Mandelbrot[9] en que genera sèries temporals que presenten el que avui en

[7] Vegi's com exemple d'aplicació Kolmogorov (1941a).

[8] En sèries de periodicitat inferior a la anual, aquests salts són també habituals a les freqüències estacionals.

[9] Vegi's Mandelbrot i Wallis (1969a) i Mandelbrot (1975b).

dia entenem com *memòria llarga*, per intentar reproduir el comportament característic de les dades temporals econòmiques.

La metodologia basada en el estadístic R/S ha estat aplicada també a dades financeres per Greene i Fielitz (1977 i 1979), Booth *et al.* (1982), Helms *et al.* (1984) i Karen i Rosenman (1986). Haubrich (1989), Haubrich i Lo (1989) i Lo (1991) usen una extensió semiparamètrica de R/S per contrastar la presència de correlació forta en el cicle econòmic, consum i evolució de preus. Posteriorment a Ooms (1994), s'ha mostrat que aquest tipus d'anàlisi no és gaire adequat per al tractament de sèries amb estacionalitat.

Tot i els treballs ja citats, el concepte de *memòria llarga* no és gaire habitual a l'Econometria tot i la presència de correlogrames amb tanta *persistència* i és manté com a metodologia habitual en aquests casos la utilització de l'operador diferenciació. Això és degut, probablement, a la no disponibilitat d'una família explícita de models en el domini temporal que presentin *memòria llarga*, mentre que els models ARMA i ARIMA són eines d'utilització molt senzilla.

A començaments de la dècada dels vuitanta es produeix un canvi radical en aquest escenari ja que Granger i Joyeaux (1980) i Hosking (1981) proposen de manera simultània i independent la utilització de l'operador diferenciació fraccional $(1-L)^d$ amb $0<d<1$ com alternativa a la diferenciació habitual. Així, aplicant l'invers d'aquest operador a un procés ε_t que sigui soroll blanc s'obté un procés X_t, que compleix la següent equació estocàstica $(1-L)^d X_t=\varepsilon_t$ i es coneix amb el nom de soroll fraccionalment integrat. Doncs bé, la principal innovació d'aquest procés és que presenta unes autocovariàncies que decreixen molt lentament i de forma hiperbòlica i que asimptòticament compleixen $|\gamma_k| \sim k^{2d-1}$. Aquest decreixement lent és la principal diferència amb els models de *memòria* limitada com poden ser els models ARMA.

A Granger (1980) es justifica la utilització de l'operador diferenciació fraccional en economia per l'efecte de l'agregació en les variables macroeconòmiques, tot i que prèviament, ja a Robinson (1978) s'hi troba una justificació a la presència de *memòria llarga* per efecte de l'agregació. Concretament, a Granger (1980) es planteja un model en el que un conjunt de variables microeconòmiques es poden representar a través d'un model AR(1) estable cada una d'elles però on els paràmetres de cada un dels models poden ser diferents. Com exemple considera el cas en que aquests paràmetres són aleatoris distribuïts segons una llei *Beta*, i demostra que en aquest cas, al agregar les variables s'obté una variable que presenta *memòria llarga*. Aquesta línia de treball l'han continuat Lin (1991) i Granger i Ding (1996).

Un cop proposat l'operador diferenciació fraccional, es desenvolupen extensions naturals d'aquest a base de combinar-lo amb polinomis autoregressius o mitjana mòbil generant així el que es coneix sota el nom de models ARFIMA (autoregressiu, fraccionalment integrat i mitjana mòbil).

Brockwell i Davis (1987) demostren que aquest tipus de models són convergents en mitjana quadràtica. Sowell (1990) proposa una estimació conjunta i màxim versemblant de tots els paràmetres del model. A partir d'aquí han aparegut tot un conjunt de propostes per a l'estimació d'aquests models.

Gray *et al.* (1989) suggereixen una extensió del model ARIMA sota el nom de model GARMA que de fet és més general que els models ARFIMA i es basa en diferències fraccionals en altres freqüències a través de l'operador $(1-2\xi L+L^2)^\lambda$. En aquest model la freqüència $\omega_0=arcos(\xi)$ presenta *memòria llarga*. Chung (1993) analitza l'estimació del paràmetre ξ però són models que presenten força dificultats tècniques pel càlcul de les autocovariàncies.

En una altra línia de treball paral·lela als models ARFIMA, Beran (1993) generalitza el model exponencial de Bloomfield (1973) i defineix un model exponencial fraccional (FEXP) que també presenta *memòria llarga*. En aquest cas, aplica l'estimació generalitzada a una regressió en el domini de les freqüències i utilitza les tècniques inferencials habituals.

A partir de totes aquestes aportacions teòriques, la literatura sobre l'aplicació dels models de *memòria llarga* a l'econometria i l'estadística augmenta molt ràpidament. Es poden trobar recopilacions d'aquestes aplicacions a Beran (1992), Baillie (1994) o Robinson (1994b), i estan apareixent durant els anys noranta importants aportacions pel que fa a l'aplicació d'aquests models al mon econòmic.

De totes maneres, tot i la ràpida acceptació de la importància del concepte, l'estimació del paràmetre fraccional que mesura el grau de *memòria* o *persistència* en aquests models és difícil, tal i com ja es destacava en el treball de Lawrance i Kottegoda (1977) i això ha retardat bastant l'aplicació d'aquest tipus de models.

Un salt endavant important es produeix arrel del treball de Geweke i Porter-Hudak (1983) que proposen un estimador semiparamètric molt senzill de l'ordre d'integració i l'usen per realitzar prediccions de sèries de preus. Ràpidament apareixen treballs que apliquen tècniques similars a aquestes amb altres finalitats, per exemple a Diebold i Rudebusch (1989) i Shea (1989 i 1991).

Totes les referències anteriors es centren en la *memòria llarga* des d'una perspectiva bàsicament estadística. Més tard apareixen altres treballs amb un punt de vista més economètric com Gourieroux *et al.* (1989), Sowell (1989 i 1990), Robinson (1994). De fet Gourieroux *et al.* (1989) i Sowell (1989) demostren que les distribucions asimptòtiques dels estimadors en models de regressió són molt diferents segons el tipus de *memòria* de les dades. A més, Yajima (1989) ha mostrat algunes de les conseqüències de no tenir en compte la presència de correlació forta per a la realització d'inferència.

Una altra font important de literatura sobre el concepte de *memòria* està relacionada amb la mesura del grau de *persistència* dels *shocks*. Cal citar en aquest aspecte Huizinga (1986), Campbelll i Mankiw (1987 i 1988), Poterba i Summers (1987), Cochrane (1988), Fama i French (1988) i Lo i MacKinlay (1988).

Totes les referències anteriors fan esment al concepte de *memòria llarga* en sèries de periodicitat anual o, com a mínim, a sèries sense estacionalitat. Pel que fa al cas de sèries estacionals, els primers models de *memòria llarga* a freqüències estacionals són suggerits per Abrahams i Dempster (1979) que es basen en el model proposat per Mandelbrot i el generalitzen per permetre la presència d'una component estacional. Aquest model té espectre infinit sols a la freqüència zero i no permet recollir la presència d'una estacionalitat infinita, però Jonas (1993) ha generalitzat aquest model per permetre valors no acotats de l'espectre a les freqüències estacionals.

Carlin *et al* (1985) intenten combinar soroll gaussià fraccional no estacional i estacional amb un model de components no observables per sèries temporals econòmiques. Degut als problemes tècnics per trobar les autocovariàncies d'aquest tipus de models aquests autors no foren capaços de trobar estimadors màxim versemblants[10].

En els darrers anys s'ha produït una molt àmplia recerca en la generalització dels models ARIMA estacionals proposats per Box i Jenkins (1970) al cas de la integració fraccional. Porter-Hudak (1990) aplica una generalització estacional a agregats monetaris de EUA utilitzant un mètode d'estimació molt senzill de regressió a partir del periodograma. Ray (1991) generalitza molts dels resultats estadístics estàndard del models fraccionalment integrats al cas estacional en que apareix l'operador $(1-L^s)^d$.

Els models fraccionalment integrats periòdics, és a dir, models on el paràmetre de la integració fraccional varia d'estació en estació són una alternativa molt interessant per a les sèries estacionals. Una aplicació d'aquests models es pot trobar a Franses i Ooms (1995) però encara està poc desenvolupada la teoria sobre la distribució dels estimadors i l'anàlisi de l'espectre i funcions d'autocovariàncies que en aquest cas són dependents del temps.

Cal destacar també que pel que fa als mètodes de diagnòstic sobre els resultats de l'estimació d'aquests models encara no està gaire desenvolupat. Recents estudis han presentat evidència de que les sèries de dades financieres es poden aproximar millor amb models integrats fraccionalment que per models amb arrels unitàries. En aquest sentit, es poden citar Diebold *et al.*, Shea (1991), Backus i Zing (1993) i Cheung (1993).

[10] Veure Carlin i Dempster (1989).

Algunes sèries econòmiques per a les que s'han estimat models de *memòria llarga* són sèries de Producte Nacional Brut (Diebold i Rudebush, 1989 i Haubrich i Lo, 1989), tipus d'interès (Shea, 1990), índexs de preus (Geweke i Porter-Hudak, 1983 i Tieslau, 1992), taxes d'inflació (Baillie, Chung i Tieslau, 1995 i Hassler i Wolters, 1995), sèries financeres (Greene i Fielitz, 1977; Helms *et al.*, 1984; Aydogan i Booth, 1988 i Lo, 1991), i tipus de canvi (Cheung, 1993 i Booth *et al.*, 1982).

A l'actualitat s'està desenvolupant una molt àmplia recerca al voltant del concepte de *memòria llarga*, tant pel que fa a mètodes d'estimació tant paramètrics com semiparamètrics que superin les limitacions dels mètodes *clàssics*, com a l'aplicació de models amb *memòria llarga* per revisar algunes conclusions prèvies sobre el comportament de certes dades econòmiques, especialment pel que fa al grau d'incidència futura que poden tenir *shocks* presents en aquestes sèries.

Entre aquests treballs més recents es poden destacar els de Porter-Hudak (1990), Sowell (1992), Linden (1995), Patterson i Sowell (1996), Chambers (1996), Hassler (1996), Hidalgo i Robinson (1996), Koustas i Veloce (1996), Lee i Robinson (1996), Cappuccio i Lubian (1997), Dolado i Marmol (1997), Gil-Alaña i Robinson (1997) o Lee i Amsler (1997).

1.3. *Memòria* versus integrabilitat

Tal com ja s'ha comentat abans, la resposta habitual en el món de les sèries econòmiques a la presència d'una gran *persistència* en les dades i per tant a la impossibilitat d'usar directament models ARMA ha estat el supòsit que les dades són no estacionàries però que a través de l'operador diferenciació és possible aconseguir aquesta estacionarietat.

Sens dubte, la popularitat d'aquest operador diferenciació entre els econòmetres no és aliena al desenvolupament de la teoria de la cointegració. Naturalment, la utilització de l'operador diferenciació en dades econòmiques és històricament previ al desenvolupament de la teoria de la cointegració tot i que és molt menys popular que en altres disciplines científiques, fins que el concepte de regressió espúria i el desenvolupament de la teoria de la cointegració acaben d'aportar eines per la seva utilització.

Com és ben conegut, la presència de no estacionarietat invalida la majoria de resultats clàssics sobre la inferència en models economètrics però sota el supòsit que les dades són integrades, aquesta teoria permet realitzar inferència estadística sobre les relacions entre variables no estacionàries. Sobre aquesta teoria s'ha generat una àmplia literatura i un gran nombre de resultats d'interès durant els darrers quinze anys entre els que cal destacar, Engle i Granger (1987), Stock (1987) i Johansen (1988).

S'ha generat així, a l'anàlisi econòmica de dades temporals una dicotomia en que es distingeixen dos possibilitats. En el primer cas, quan les dades són estacionàries, és normal suposar que, al marge de components deterministes que prèviament cal detectar i separar, les observacions es poden caracteritzar com una realització d'un procés ARMA(p,q) amb ordres p i q no gaire grans[11]. El treball de Box i Jenkins fou fonamental pel desenvolupament d'unes eines que permeten la identificació i estimació del model d'aquest tipus més adequat a les dades.

En canvi, quan les dades no tenen un comportament estacionari és habitual recórrer a la diferenciació per transformar-les en estacionàries. Sorgeixen així els models ARIMA com extensions naturals dels models ARMA. Com s'ha mostrat en gran nombre de treballs a la dècada dels anys vuitanta, en el context de les sèries integrades no diferenciades, es presenta la possibilitat d'obtenir regressions espúries. Així, l'estimació de relacions economètriques entre variables integrades passa per la utilització de la teoria de la cointegració.

[11] En el cas de sèries de periodicitat inferior a l'anual, calen models estacionals ARMA(p,q)(P,Q)$_s$.

Les grans diferències que hi ha en el tractament d'ambdós tipus de situacions han generat que a l'actualitat, l'anàlisi economètrica de dades temporals sigui molt depenent dels tests usats per discriminar entre una o altra situació, o tests d'arrels unitàries.

Deixant de banda les dificultats que poden tenir aquests tests pel que fa a la seva aplicació pràctica amb tamanys mostrals reduïts, bona part del tractament actual de les dades econòmiques temporals està basada en la dicotomia entre models ARMA i models ARIMA. Però el fet d'establir una classificació de les dades temporals entre models ARMA i models ARIMA presenta, almenys, dos problemes:

a) D'una banda, hi ha certa unanimitat a la literatura en que els models integrats no són més que una modelització molt particular de la presència de la no estacionarietat. De totes maneres, aquest és un problema que té molt difícil solució ja que no és fàcil el desenvolupament d'eines en un entorn no estacionari de manera que com a única solució, de moment, queda la de restringir l'estudi de la no estacionarietat a l'anàlisi de certes transformacions, com la diferenciació, que porten a que les dades transformades siguin estacionàries.

b) L'altra problema pel que creiem que no s'ha mostrat prou interès fins al moment i que en certa manera inspira la realització d'aquesta tesi, és la no equivalència entre la condició d'estacionarietat i l'existència d'una representació en forma de model ARMA(p,q). En aquest sentit, l'aplicació de la metodologia proposada per Box i Jenkins a les dades econòmiques es basa en la hipòtesi que amb ordres p i q relativament petits es poden aproximar força bé les dades observades[12].

Però com ja s'ha comentat els models ARMA presenten unes autocorrelacions que decreixen de manera exponencial i en principi, no hi ha cap raó per restringir l'anàlisi a models amb aquesta característica. De fet, que l'autocorrelació dels models ARMA sigui altament

[12] Ja en els treballs de Whittle (1951) i Hannan (1970, pàg. 245) es destaca que, de vegades, són necessaris un gran nombre de paràmetres per ajustar un model ARMA adequat.

decreixent és un dels instruments bàsics usats a la pràctica per detectar si una sèrie és o no estacionària ja que s'associa un decreixement més lent de la funció d'autocorrelació a la no estacionarietat. Però una funció d'autocorrelació que decreix lentament no és incompatible amb l'estacionarietat. És clar que una condició necessària per a l'estacionarietat és que les autocorrelacions tendeixin a zero, mentre que els models ARMA compleixen una condició més forta, concretament que la suma en valor absolut d'aquestes autocorrelacions és finita.

Però a nivell conceptual, si no es restringeix l'anàlisi a cert tipus de models, una sèrie temporal pot ser estacionària i que la suma en valor absolut de les autocorrelacions sigui infinita. En altres paraules, és possible que una sèrie temporal tot i ser estacionària presenti una major dependència dels valors passats que la que presenten les sèries generades per models ARMA.

Precisament, el concepte de *memòria llarga* està associat, de manera intuïtiva, a aquesta major dependència de valors passats, anomenada de vegades com *persistència* o com *memòria*. Fins al moment, els diferents models teòrics usats per representar aquesta *memòria llarga* presenten un decreixement de les autocorrelacions de tipus hiperbòlic, però en principi el comportament podria ser més complicat, sols cal, per a que una sèrie sigui estacionària que l'àrea per sota de l'espectre sigui finita.

Per tant, la divisió de les sèries temporals en sèries I(0) modelades com ARMA i sèries I(1) és insuficient, perquè és possible trobar-se amb sèries I(0) que no és correcte modelitzar-la a través de models ARMA.

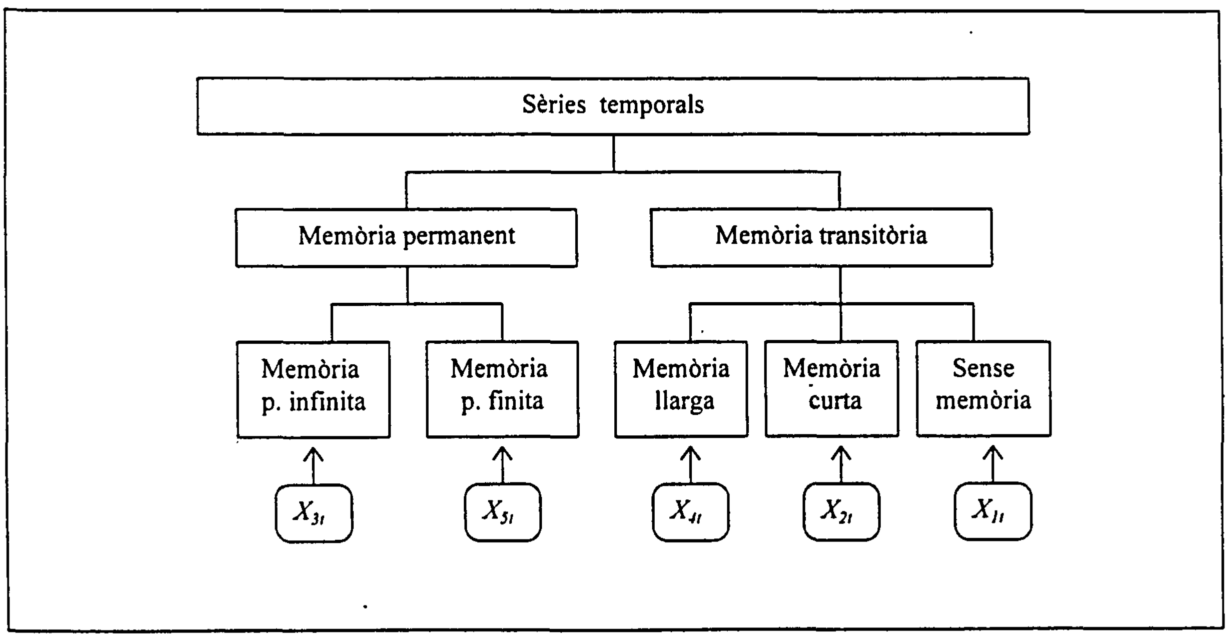
A més, entre les sèries estacionàries amb *memòria llarga* i les sèries I(1) encara hi ha una altra possibilitat. Les sèries I(1) a més de ser no estacionàries, tenen la propietat que qualsevol shock que afecti a la variable té un efecte permanent en les observacions futures de la variable. Però és possible, almenys a nivell conceptual que una sèrie sigui no estacionària amb una *memòria permanent* però que els shocks no tinguin efectes permanents.

Per tant, la diversitat de situacions és molt més complexa que la que es dedueix de la dicotomia entre variables I(0) i variables I(1). Així, a mode d'exemple, consideri's 5 sèries temporals amb les següents característiques:

- Sèrie X_{1t} soroll blanc, és a dir, $E(X_{1t})=0$ i $E(X^2_{1t})=\sigma^2$.
- Sèrie $X_{2t} \sim$ ARMA(p,q) i, per tant $X_{2t} \sim$I(0).
- Sèrie $X_{3t} \sim$ ARIMA(p,1,q) i, per tant $X_{3t} \sim$I(1).
- Sèrie X_{4t} estacionària però amb una successió de coeficients d'autocorrelació que no compleixen la condició de sumabilitat absoluta.
- Sèrie X_{5t} no estacionària però amb reversió a la mitjana.

Sense entrar en detalls que s'analitzaran més formalment al capítol 3, es poden relacionar les característiques de cada una d'aquestes sèries amb les diferents categories del quadre 0.1. de la Introducció. Així, al quadre 1.1 s'han distribuït de manera adient cada una de les 5 sèries per mostrar com la distinció entre sèries I(0) i sèries I(1) és una simplificació extrema d'una diversitat molt més complexa.

Quadre 1.1. Exemples de sèries temporals amb diferents tipus de *memòria.*



1.4. Els models ARFIMA i els diferents tipus de *memòria*

Atesa la no coincidència dels conceptes de *memòria* i d'integrabilitat, cal disposar en primer lloc d'una definició precisa dels diferents tipus de *memòria* que permeti una divisió de les sèries temporals segons aquesta nova característica. Però a més, per a que aquesta nova divisió sigui aplicable a efectes pràctics és necessari disposar d'algun tipus de modelització explícita de les diferents formes de *memòria*. De fet, al nostre entendre, la no disponibilitat d'aquest tipus de model és la causa que durant molts anys l'anàlisi aplicada s'hagi continuat basant en la divisió de les sèries en I(0) i sèries I(1).

Per sort, es disposa actualment d'una modelització explícita molt còmode per tractar amb els diferents tipus de *memòria*. Es tracta dels models coneguts com ARFIMA. Recordi's que un model ARIMA(p,d,q) es pot representar com:

$$(1-\phi_1-...-\phi_p L^p)(1-L)^d X_t = (1-\theta_1 L-...-\theta_q L^q L)\varepsilon_t \tag{1.1}$$

on ε_t és soroll blanc, l'ordre d'integrabilitat d és zero o un valor natural, i els polinomis $\phi(L)=1-\phi_1-...-\phi_p L^p$ i $\theta(L)=1-\theta_1-...-\theta_q L^q$ tenen totes les arrels fora del cercle unitat.

A partir d'aquest model, la definició dels models ARFIMA(p,d,q) és ben senzilla: consisteix a permetre que l'ordre d'integrabilitat pugui ser qualsevol nombre racional. De fet, n'hi ha prou amb que es compleixi la condició $-1/2<d<1/2$ per garantir que tot procés X_t que compleixi l'equació (1.1) sigui estacionari.

La utilitat del model (1.1) és que inclou com a casos particulars gran varietat de situacions[13]:

- Si $d=0$, no hi ha comportament de llarg termini, sinó que la sèrie sols presenta *memòria curta* i el model (1.1) es transforma en un model ARMA.

[13] Al capítol 4 s'hi pot trobar una anàlisi més formal dels models ARFIMA.

- Si $d=1$, es tracta d'un model ARIMA de manera que X_t és no estacionari i els shocks que afecten a X_t tenen efectes permanents.

- Si $0<d<1/2$ es pot comprovar que X_t és una sèrie estacionària amb una funció d'autocorrelació que decreix molt lentament, tant que l'espectre presenta una singularitat en la freqüència zero. Però tot i que hi ha una singularitat en tots els casos, diferents valors de d donen lloc a diferents graus de *persistència*. Aquesta és la traducció formal de l'evidència empírica de que moltes sèries econòmiques, tot i tenir comportaments ben diferents, semblen iguals en el domini de les freqüències.

- Si $1/2 \leq d<1$, la sèrie presenta unes autocorrelacions que decreixen molt lentament, però ara no és estacionària. La novetat d'aquest cas és que a diferència del cas $d=1$, la sèrie presenta reversió a la mitjana, de manera que els shocks no tenen efectes permanents. És a dir, un shock que afecti al valor de la variable en un moment donat, pot tenir efectes durant molt temps (molt més que en un model ARMA que té una *memòria* molt limitada), però aquest efecte no és permanent en el temps a diferència del que passa quan $d=1$.

Així, sense entrar en detalls formals, es pot comprovar com aquesta modelització permet una major flexibilitat en el moment d'estudiar el grau de *persistència* de les variables econòmiques i en la seva modelització que la restricció a models ARMA o ARIMA.

Quadre 1.2. Característiques model ARFIMA(p,d,q) segons el valor de d.

d	*Memòria*	Estacionarietat	Model
(-1/2,0)	Transitòria / Curta	Si	ARFIMA
0	Transitòria / Curta	Si	ARMA
(0,1/2)	Transitòria / Llarga	Si	ARFIMA
1/2	Permanent / Finita	No	ARFIMA
(1/2,1)	Permanent / Finita	No	ARFIMA
1	Permanent / Infinita	No	ARIMA
(1,∞)	Permanent / Infinita	No	ARFIMA

A més, tal i com es demostra al capítol 4 de la tesi, hi ha una relació directa entre el valor del paràmetre d d'un model ARFIMA i la classificació presentada anteriorment. Així, avançant-nos a resultats posteriors el quadre 1.2 mostra, de manera esquemàtica, aquestes relacions.

Una forma alternativa d'interpretar l'existència de comportaments que es troben a mig camí entre el comportament de models ARMA i models ARIMA es troba en l'evidència empírica que sovint s'accepta la necessitat d'usar l'operador diferenciació perquè s'entén que les dades presenten un comportament no estacionari mentre que les dades diferenciades mostren en el domini de les freqüències símptomes de diferenciació. De fet, la majoria de sèries econòmiques diferenciades presenten densitats espectrals estimades properes al zero en les freqüències més baixes, fet que és sinònim de sobrediferenciació, tot i que els tests d'arrels unitàries basats en aquesta característica presenten bastants problemes amb tamanys mostrals reduïts.

A més, el fet de suposar una arrel unitària, implica acceptar que la sèrie analitzada és no estacionària i que els shocks tenen un efecte permanent. Així, el fet d'haver d'escollir entre si una sèrie és I(0) o és I(1) té unes conseqüències molt importants, sense que molt cops sigui fàcil decidir-se per alguna de les dues opcions. A més, l'especificació en forma de model AR(1) no és l'única que permet contrastar la hipòtesi d'una arrel unitària. De fet, usar AR(1) té la desagradable propietat que hi ha una discontinuïtat en les distibucions límit. Hi ha altres descripcions matemàtiques que permeten contrastar la presència d'una arrel unitària, però potser la més atractiva, ja que manté un nivell de continuïtat en el seu comportament major que en els models AR(1), són els models integrats fraccionalment.

Aquests models són, amb gran diferència, els models més usats a l'anàlisi economètrica per modelitzar la presència de *memòria llarga*. Però, a diferència del que passa amb la *memòria curta* en que n'hi ha prou amb considerar la classe dels models ARMA, en el cas de la *memòria llarga* sols es disposa d'alguns models concrets però no d'arguments teòrics que justifiquin aquests o altres models, de moment desconeguts.

Probablement, la seva popularitat recau en el fet que són una extensió bastant natural dels models ARIMA clàssics i per tant, són més propers als economètres i economistes. Aquesta popularitat fa que en alguna ocasió es pugui arribar a pensar que hi ha una equivalència entre el concepte de *memòria llarga* i models ARFIMA però això no és cert. Així, tot i que al llarg de la tesi es fa referència bàsicament a aquest tipus de models, no cal oblidar que no són més que un tipus de model molt concret però, de fet, gairebé l'únic que es disposa.

Entre els inconvenients que presenten els models ARFIMA n'hi ha dos especialment importants que cal destacar:

a) Un model ARFIMA no és més que la representació amb un número reduït de paràmetres d'un ARMA amb ordres p i/o q infinits. Per tant, en el cas estacionari, el teorema de Wold garanteix que es pot aproximar el seu comportament tant bé com es vulgui usant models ARMA amb ordres finits. Com que la identificació, estimació i diagnòstic de models ARMA és un problema prou conegut i resolt de manera satisfactòria, no està clar quins avantatges aporta la utilització d'aquest tipus de models[14].

b) Un paràmetre d natural té una interpretació clara en termes econòmics, però no és així quan d és un número racional.

Donades les complicacions metodològiques de la consideració de la presència de *memòria llarga* en relació a la simplicitat analítica i interpretativa dels models ARMA i ARIMA, cal plantejar-se seriosament les avantatges de la consideració d'aquests models. De fet, un dels objectius de la tesi és clarificar en quines situacions és rellevant la consideració d'aquestes noves tècniques.

[14] Models que d'altra banda, i com s'analitza al llarg de la tesi presenten molts més problemes pel que fa a la seva identificació, estimació i utilització que els models ARMA.

De fet, la justificació dels nous models no pot basar-se sols en la constatació que la utilització de les tècniques habituals presenta problemes sinó que cal provar que els nous models permeten millorar els resultats obtinguts.

1.5. La rellevància economètrica del tipus de *memòria*

De fet, amb diferents matisos el concepte de *memòria* és present a bona part del desenvolupament de les eines per l'anàlisi de sèries temporals econòmiques:

1) D'una banda, des de un punt de vista més estadístic, la noció de *memòria* està associada al conjunt de restriccions que cal imposar a una sèrie temporal per garantir propietats dels estimadors. Així, està associada al concepte d'ergodicitat, o a conceptes com els de mixtura forta (Rosenblatt, 1956) i mixtura uniforme (Ibraginov, 1959).

2) D'altra banda, des d'un punt de vista més economètric, s'ha mostrat molt interès per la persistència dels *shocks* econòmics. Treballs com els de Nelson i Plosser (1982), Cochrane (1986), Campbelll i Mankiw (1987) o Fama i French (1988), entre altres, mostren el seu interès pel fet que tot i que moltes sèries, per exemple la del PIB, semblen mostrar una dinàmica a curt termini diferent d'un camí aleatori amb deriva, els shocks són molt persistents en el temps. De fet, la polèmica sobre si els shocks són o no permanents està molt relacionada amb la discriminació entre tendències estocàstiques i tendències deterministes.

Més concretament, si el PIB en termes reals conté una arrel unitària o tendència estocàstica això recolza la hipòtesi dels cicles econòmics reals ja que està generalment acceptat que canvis permanents en el nivell del PIB només es poden interpretar com conseqüència de canvis en la productivitat. En el context de tendències estocàstiques, qualsevol shock sobre el sistema econòmic tindrà efectes permanents, de manera que cal prendre alguna mesura de política econòmica per a que la variable torni a la trajectòria original de llarg termini. D'altra banda, si s'accepta la presència d'una tendència determinista enlloc d'una tendència estocàstica, les fluctuacions són transitòries i per tant hi ha menys necessitat de prendre mesures polítiques.

La discriminació entre diferents tipus de *memòria* és també important en molts altres camps de l'anàlisi econòmica basats no en les dades originals sinó amb les dades que s'obtenen un cop s'hi ha extret la tendència. Òbviament, el mètode adequat per a l'extracció d'aquesta tendència és diferent segons de quin tipus sigui aquesta, de manera que el diagnòstic adequat del tipus de tendència de les sèries econòmiques és molt important. En aquest sentit, és ben conegut que l'eliminació de tendències deterministes quan les dades s'han generat amb una d'estocàstica produeix un comportament cíclic espuri, mentre que si s'utilitza l'operador diferenciació quan és de tipus deterministes la sèrie diferenciada presenta variacions espúries a les altes freqüències.

Ja a Granger (1966) es destaca que la majoria de sèries econòmiques presenten un predomini absolut de les baixes freqüències, equivalent a que hi ha una forta persistència de les autocorrelacions. Aquest autor destaca que l'espectre estimat (o fins i tot el periodograma) de moltes dades econòmiques suggereix que la densitat espectral pugui ser infinit a la freqüència zero (en dades amb forta estacionalitat també es presenta aquesta característica en les freqüències estacionals).

En el treball de Box i Jenkins, es proposa salvar aquesta dificultat amb l'ampliació dels models ARMA per la via d'usar l'operador diferències, és a dir, usant models ARIMA. També ha estat fonamental la teoria de la cointegració que permet l'anàlisi de relacions economètriques entre variables integrades quan no són aplicables les tècniques clàssiques d'inferència estadística.

Però les arrels unitàries, incloent múltiples arrels o arrels complexes del cercle unitat, són una part molt petita de les possibles descripcions matemàtiques de la no estacionarietat. De vegades, l'existència d'una arrel unitària té un clar significat econòmic, però molta de la popularitat de les arrels unitàries prové de que sovint no són aplicables els models no estacionaris alternatius, i de la simplicitat del plantejament de les arrels unitàries. A Sims (1988) s'hi pot trobar una forta crítica a alguns aspectes de l'interès economètric actual per les arrels unitàries.

La característica fonamental que fa útil els models integrats per superar les limitacions dels models ARMA és que s'aconsegueix modelitzar la presència de singularitats a l'espectre d'una sèrie temporal. La teoria de la integració i de la cointegració han permès tractar de manera separada el comportament a curt termini del comportament a llarg termini. Però la hipòtesi de l'arrel unitària potser és massa restrictiva. D'una banda, hi ha un salt conceptual important entre el cas I(0) i el cas I(1), de manera que els tests que permeten contrastar una hipòtesi contra l'altra sovint tenen una baixa potència.

D'altra banda, l'evidència empírica de les sèries econòmiques revela que romanen problemes amb aquesta metodologia. En particular, hi ha moltes sèries que presenten una funció d'autocorrelació amb un decreixement molt lent, però que un cop diferenciades presenten un espectre nul a la freqüència zero, símptoma que s'ha sobrediferenciat. Ja Engle (1976) recomana complementar els estadístics usats en l'especificació de models ARIMA amb diagnòstics a través de l'estimació no paramètrica de l'espectre que té l'avantatge de no basar-se en cap especificació paramètrica.

Per adonar-se'n de la importància del concepte de *memòria* pel que fa a l'anàlisi de sèries temporals econòmiques n'hi ha prou amb considerar un exemple. Suposi's que es genera una successió de variables aleatòries U_t, independents i idènticament distribuïdes segons una distribució uniforme a l'interval $[0,1]$. Si Z és una variable aleatòria normal estàndard i es defineix un procés estocàstic com $X_t=U_t+Z$ es pot comprovar que X_t és estrictament estacionari però la mitjana mostral:

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t \qquad (1.2)$$

no convergeix cap a $E[X_t]=1/2$. De fet, es compleix que [15]:

$$\overline{X}_T \xrightarrow{q.s.} \frac{1}{2} + Z \qquad (1.3)$$

El problema en aquest cas, és que hi ha massa dependència en la successió X_t. Per garantir bones propietats estadístiques dels estimadors habituals a l'anàlisi de sèries temporals, cal imposar alguna restricció en la *memòria* del procés. La més habitual és la noció d'ergodicitat[16].

Tot i que es pot definir el concepte d'ergodicitat en un context més general, per la nostra argumentació n'hi ha prou a considerar la definició més senzilla restringida al cas d'estacionarietat[17].

Però tot i que un procés sigui ergòdic pot haver-hi també problemes. Per exemple, tot i que la mitjana mostral sigui consistent, pot ser-ho a un rati molt inferior a l'habitual. Per comprovar-ho, consideri's el següent exemple:

La mitjana mostral d'una sèrie $X_1,..,X_T$ estacionària té per esperança μ i per variància:

$$Var(\overline{X}) = E(\overline{X}-\mu) = T^{-2}E\left[\sum_{t=1}^{T}(X_t-\mu)\right]^2 = T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\gamma_{t-s} = T^{-2}\left\{2\sum_{t=1}^{T}(T-t)\gamma_t + T\gamma_0\right\} \qquad (1.4)$$

on γ_k són les autocovariàncies d'X_t.

Interessa conèixer les característiques asimptòtiques d'aquesta mitjana mostral. Observi's que si el límit de $TVar(\overline{X})$ quan T tendeix cap a infinit és finit, el seu càlcul és ben senzill.

$$\lim_{T\to\infty} TE(\overline{X}-\mu)^2 = \sum_{j=-\infty}^{\infty}\gamma_j = f(0) \qquad (1.5)$$

[15] Es tracta de convergència quasi segura que per tant implica convergència en probabilitat, la més habitual en l'anàlisi economètrica.

[16] Tot i que recentment s'han desenvolupat altres teories més generals, com les basades en mixtures.

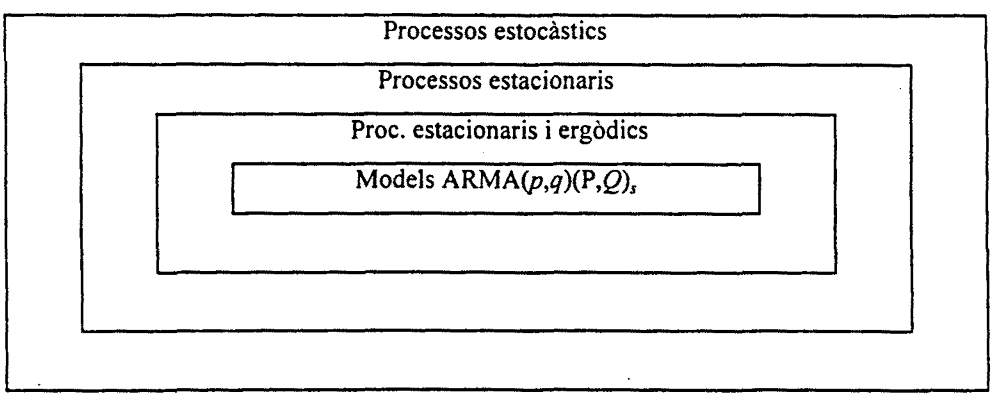
[17] A efectes d'aquesta exposició, un procés estacionari és ergòdic si els moments mostrals convergeixen en probabilitat als moments poblacionals.

on $f(\cdot)$ és la densitat espectral de la sèrie. Però pot ser que no existeixi el límit de $TVar(\bar{X})$ o bé que no sigui finit. Molts dels teoremes asimptòtics referits a sèries temporals imposen la restricció de la sumabilitat absoluta de les autocovariàncies, és a dir, $\sum|\gamma_j| < \infty$. Aquesta condició s'utilitza habitualment com una condició suficient d'ergodicitat.

Observi's que, en particular, aquesta condició implica que $f(0) = \sum \gamma_j < \infty$ [18] de manera que si les autocovariàncies són absolutament sumables, llavors és immediat comprovar, a partir de (1.5) que la mitjana mostral convergeix cap a l'esperança del procés amb ordre de convergència d'ordre $O(T^{1})$.

Però la condició de sumabilitat absoluta no és equivalent a l'ergodicitat, de manera que s'ha d'acceptar la possibilitat que una sèrie temporal sigui ergòdica però les autocovariàncies decreixin prou lentament per a que no siguin absolutament sumables. En aquest cas, el sumatori de la part dreta de l'expressió (1.4) tendeix cap a zero molt més lentament i per tant la mitjana mostral convergeix més lentament cap a l'esperança del procés.

Quadre 1.3. Relació entre estacionarietat, ergodicitat i models ARMA.



De fet, al llarg de la tesi s'il·lustra com una sèrie estacionària i ergòdica pot presentar unes autocovariàncies que decreixin prou lentament, per a que la mitjana aritmètica, tot i ser un

[18] Pel teorema de convergència dominada, la condició de sumabilitat absoluta permet demostrar l'existència del límit anterior.

estimador consistent de la seva esperança, sigui un estimador molt dolent quan es disposa de tamanys mostrals que són habituals a l'anàlisi de dades econòmiques. Això és precisament el que passa sota *memòria llarga.*

Aquest exemple tan senzill mostra com, fins i tot en el cas d'un estadístic tan habitual com la mitjana mostral, hom es pot trobar amb una gran diversitat segons el grau de *memòria* que presentin les dades temporals.

En tot cas, l'aspecte fonamental que fa necessari considerar el concepte de *memòria llarga* és la transcendència que té la presència d'aquest fenomen per a la inferència economètrica amb dades temporals. Evidentment, no té sentit incloure un conjunt nou de models, que a més presenten bastants problemes de tipus metodològic, únicament perquè són una ampliació. Seguint el principi de parsimònia, sols té sentit la utilització d'aquests models si el fet de no considerar-los té conseqüències greus per a l'anàlisi economètrica.

Pel que fa a aquestes conseqüències, la presència de *memòria llarga* en les dades econòmiques pot tenir conseqüències cara a la no consideració d'aquest fenomen que es poden agrupar en dos grups:

- Pel que fa a l'anàlisi de tipus univariant de les dades, usar una representació incorrecta de les dades pot afectar a la qualitat de les prediccions. En aquest cas, l'anàlisi de les conseqüències només pot estar basat en simulacions o anàlisis teòriques en que es suposa conegut el vertader model. Però a la pràctica tant els ARIMA com els ARFIMA no són més que casos molt particulars de parametritzacions, de manera que probablement cap de les dues sigui la vertadera. En aquesta situació no està clar quina és la millor estratègia, si restringir l'anàlisi a uns models més senzills com són els models ARIMA o bé ampliar l'espectre de models considerat amb el cost d'augmentar la seva complexitat.

- Un altre tipus de conseqüència que considerem pot ser més greu és la que pot tenir en el moment de realitzar inferència economètrica el fet de no tenir en compte el tipus de *memòria* de les dades usades. Així, com a punt de partida es pren el cas més simple, és a dir, un model de regressió lineal simple:

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

En aquest context s'analitzen les conseqüències que té la presència dels diferents tipus de *memòria* en la variable explicativa i/o el terme de pertorbació pel que fa a les propietats dels estimadors habituals.

Com el tractament del model anterior és diferent segons si es considera que les variables són estacionàries o són I(1), és habitual l'aplicació d'algun test d'arrels unitàries a les variables. Aquesta eina és fonamental per prevenir que puguin obtenir-se regressions espúries. En aquest sentit, s'ha analitzat el comportament de molts d'aquests tests i presenten bastants problemes quan els processos generats presenten *memòria llarga* o *memòria permanent finita.*

Però, al marge d'aquests problemes, fins i tot sota el supòsit que s'encerta el caràcter d'estacionarietat o no estacionarietat de les dades, el fet de no considerar el tipus de *memòria* pot portar a conclusions errònies pel que fa, entre altres, als següents aspectes:

- D'una banda, la representació a través de models paramètrics com poden ser, per exemple, els models ARMA.
- D'altra banda, la realització d'inferència sobre l'existència de relacions entre variables.

Un bon exemple d'aquests problemes n'és la freqüència amb que s'utilitzen models tipus línies aèries per modelitzar dades econòmiques, tot i que sovint s'ha argumentat sobre el perill de que en aquests casos s'estiguin sobrediferenciant les dades. Òbviament, si es produeix aquesta sobrediferenciació, les conseqüències no són, per conegudes, menys greus. Entre altres, cal destacar les següents:

- La no invertibilitat de la representació mitjana mòbil i per tant, la inconsistència dels estimadors de paràmetres autoregressius i models VAR

- La mala especificació del model i per tant, errors en l'estimació de components no observables i en la realització de prediccions.

En definitiva, de les consideracions anteriors es dedueix que el concepte de *memòria* és una característica de les sèries temporals força important i que cal tenir en compte a l'hora d'utilitzar aquestes dades com a base per algun tipus d'anàlisi economètrica. Tot i això, es tracta d'una característica que no ha estat objecte de massa atenció per part dels econòmetres fins no fa gaire, encara que últimament s'està produint molta recerca a l'Econometria al voltant del concepte de *memòria llarga*.

En tot cas, els arguments presentats, tot i que de manera informal, al llarg d'aquest capítol, demostren la necessitat d'analitzar amb cert detall la sensibilitat de les eines economètriques més habituals front a errors en l'especificació del tipus de *memòria*, tasca que s'aborda a partir del següent capítol.

Capítol 2

MESURES DE LA *PERSISTÈNCIA* EN VARIABLES ECONÒMIQUES

2.1. Introducció

Prèviament a la utilització del concepte de *memòria llarga*, en el tractament de variables econòmiques es produeixen en diferents àmbits de l'anàlisi econòmica aportacions que estan, d'alguna manera, molt associades amb les característiques que motiven el desenvolupament d'aquest concepte.

Un dels aspectes que més relació en té és la controvèrsia al voltant del tipus de tendència que presenten les dades econòmiques. En aquest sentit, una de les característiques més clares de les sèries econòmiques és l'observació empírica de que sovint, aquest tipus de dades no contenen comportaments estacionaris, sinó tendencials.

Segons el tipus de modelització usat per representar aquestes tendències, la interpretació que cal donar a les oscil·lacions al voltant d'aquestes tendències són ben diferents. Un dels casos més clars de les diferents implicacions econòmiques d'usar diferents tipus de modelitzacions el trobem en l'anàlisi de l'evolució temporal del PIB.

Aquesta variable, segons la teoria clàssica dels cicles reals, presenta una tendència a llarg termini estable, de manera que les variacions o *shocks* que es puguin produir en la variable en un moment donat no tenen efecte permanent sobre l'evolució futura de la variable. Aquesta perspectiva és plenament coherent amb la modelització de les sèries macroeconòmiques a partir de tendències deterministes.

Així, tot i l'interès generat al voltant dels models amb arrels unitàries pels treballs de Box i Jenkins (1970) i Dickey i Fuller (1979), l'aproximació a les sèries macroeconòmiques amb tendències deterministes és predominant fins al treball de Nelson i Plosser (1982) en que s'aporta evidència sobre la presència d'arrels unitàries en sèries macroeconòmiques anuals dels EUA.

A partir d'aquest treball s'ha generat una àmplia literatura amb l'objectiu de determinar quina de les dos aproximacions és millor. Simplificant molt, un grup de models molt usats per descriure el comportament de les desviacions respecte de la tendència a llarg termini està format per una tendència determinista amb shocks estacionaris:

$$X_t = \mu t + \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j} \tag{2.1}$$

on ε_t és un procés soroll blanc

Si el segon terme de (2.1) és estacionari, les fluctuacions de la variable X_t són temporals i es diu que X_t és estacionària amb tendència determinista (sovint notat com TS).

Per altra banda, el model més senzill que permet modelitzar una variable amb fluctuacions permanents és el camí aleatori amb deriva[1] (sovint notat com DS):

$$X_t = \mu + X_{t-1} + \varepsilon_t \tag{2.2}$$

La necessitat de disposar de mètodes estadístics que permetin decidir quin dels dos tipus de modelitzacions és més adequada per les dades econòmiques, ha generat una àmplia literatura al voltant dels coneguts com *tests d'arrels unitàries*.

Però aquestes dues possibilitats són dos casos extrems d'una àmplia casuística. Així, diferents autors han proposat diferents enfocaments per analitzar els casos que hi ha entre ambdós extrems i algunes d'aquestes vies es basen en mètodes de mesurar el grau de persistència dels *shocks* sobre l'evolució futura de la variable.

De fet, algunes d'aquestes mesures han estat usades també per proposar tests d'arrels unitàries alternatius als més habituals.

[1] Es tracta d'un tipus molt concret de modelització no estacionària, a diferència del model (3.1) en que un cop eliminada la tendència, les fluctuacions són estacionàries.

Hi ha també altres mesures de *persistència* originàries d'altres àmbits científics però que, d'alguna manera, són conceptualment properes al concepte de *memòria*. Totes aquestes mesures, tot i els problemes que presenten, estan molt relacionades amb els diferents tipus de *memòria* que es presenten al capítol 3. Així, es presenten a continuació algunes d'aquestes mesures amb la intenció de relacionar-les després amb la caracterització de sèries temporals segons el tipus de *memòria*.

2.2. La funció de resposta a l'impuls

En el sí de l'anàlisi de la *persistència* dels shocks econòmics s'ha usat sovint, durant la dècada dels anys vuitanta, la funció de resposta a l'impuls per mesurar aquesta persistència. La motivació és ben senzilla: tota sèrie estacionària es pot representar en forma de mitjana mòbil infinita, un cop s'han eliminat els components deterministes. Suposi's que la sèrie X_t objecte d'anàlisi es pot representar de la següent manera:

$$X_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j} \tag{2.3}$$

on ε_t és un procés soroll blanc.

A la successió de coeficients φ_j se la coneix amb el nom de funció de resposta a l'impuls del model (2.3) i de vegades, funció de memòria. Sota aquesta modelització és ben senzill mesurar la persistència dels *shocks* sobre l'evolució futura de la variable d'interès. De fet, l'efecte d'un shock en el moment *T-k* es tradueix en una variació igual a φ_k en el moment *T*.

De totes maneres, una representació com (2.3) sols és vàlida sota el supòsit que X_t és estacionari, de manera que $\sum_{j=0}^{\infty} \varphi_j^2$ és finita i, per tant, la successió de coeficients φ_j tendeix cap a zero necessariament. Això implica que la influència de possibles fluctuacions de la variable desapareix a llarg termini.

Molt relacionat amb la funció de resposta a l'impuls i després de la introducció del concepte d'arrels unitàries, molts autors usen els tests d'arrels unitàries per concloure si hi ha o no major *persistència*, però Campbell i Mankiw (1987) argumenten que aquesta presència és compatible amb diferents graus de *memòria*. Per exemple, en un cas tan senzill com el d'un model IMA(1,1), $(1-L)X_t = \mu + (1-\theta L)\varepsilon_t$ l'efecte d'un shock en el moment *T-k* es tradueix en una variació igual a $\varphi_k = 1-\theta$ en el moment *T* ja que:

$$\frac{1-\theta L}{1-L}\varepsilon_j = (1+L+L^2+..)(1-\theta L)\varepsilon_j = \sum_{j=0}^{\infty}(1-\theta)\varepsilon_j \qquad (2.4)$$

Així, alguns autors com Campbell i Mankiw (1987) o Cochrane (1986) proposen diferenciar les dades tot i el perill de sobrediferenciar perquè aquesta sobrediferenciació no afecta al supòsit d'estacionarietat. Així, es proposa usar el model següent:

$$(1-L)X_t = \mu + b(L)\varepsilon_t \qquad (2.5)$$

on, en principi, es permet al polinomi *b(L)* que contingui arrels unitàries. A partir d'aquesta igualtat es troba una representació mitjana mòbil per a la sèrie X_t en nivells:

$$X_t = \mu t + a(L)\varepsilon_t \qquad (2.6)$$

on $b(L) = (1-L)a(L)$.

Observi's que en aquestes condicions, els coeficients de la funció de resposta a l'impuls són iguals a $\varphi_k = b_k = \sum_{j=0}^{k} a_j$.

Si $b(1) = \sum_{j=0}^{\infty} b_j = 0$, es diu que les dades presenten reversió a la mitjana de manera que els shocks no tenen un efecte permanent. De fet, que *b(1)=0* significa que el polinomi *b(L)* conté una arrel unitària i per tant, que la sèrie X_t és estacionària. En tot cas, l'estadístic *b(1)* mesura quin és, en el límit, l'efecte dels shocks a mesura que passa el temps.

2.3. La component a llarg termini d'una sèrie temporal

Beveridge i Nelson (1981) introdueixen el concepte de component de llarg termini que després s'ha convertit en una tècnica molt utilitzada per a l'anàlisi de tendències estocàstiques. Donat un procés estocàstic X_t que pot ser, tant I(0) com I(1), suposi's que ΔX_t és un procés lineal[2]:

$$\Delta X_t = b(L)\varepsilon_t \tag{2.7}$$

on ε_t són independents i idènticament distribuïts amb $E(\varepsilon_t)=0$ i $E(\varepsilon_t^2)=\sigma^2_\varepsilon$. Una expressió equivalent a l'anterior és la següent:

$$X_t = a(L)\varepsilon_t \tag{2.8}$$

on $a(L)=b(L)(1-L)^{-1}$. Cal tenir en compte però, que si X_t és I(1) aquesta darrera expressió no és invertible, en canvi si X_t és I(0) el terme $(1-L)$ del numerador i denominador es cancel·len. En tot cas, un model ARMA és un cas particular de l'anterior.

Si es disposa d'informació fins al moment t, la predicció òptima, és a dir, que minimitza l'error quadràtic mitjà de ΔX_{t+k} és:

$$E(\Delta X_{t+k}|X_t, X_{t-1},..) = \sum_{i=0}^{\infty} b_{k+i}\varepsilon_{t-i} \tag{2.9}$$

Com la sèrie objecte de l'anàlisi és X_t, la predicció òptima de X_{t+k} és:

$$\begin{aligned} E(X_{t+k}|X_t, X_{t-1},..) &= E(\Delta X_{t+k} + \Delta X_{t+k-1} + \ldots + \Delta X_{t+1} + X_t | X_t, X_{t-1},..) = \\ &= X_t + \left(\sum_{i=1}^{k} b_i\right)\varepsilon_t + \left(\sum_{i=2}^{k+1} b_i\right)\varepsilon_{t-1} + \ldots \end{aligned} \tag{2.10}$$

Observi's com l'efecte d'un shock unitari en el moment t sobre la predicció de la variable X_{t+k} és:

---

[2] Sota el supòsit, sense pèrdua de generalitat de que $E(\Delta X_t)=0$.

$$B_k = \left(\sum_{i=1}^{k} b_i \right) \tag{2.11}$$

Si es considera a continuació, la predicció d'un valor futur infinitament llunyà, s'obté que la predicció és, fent tendir k cap a ∞:

$$\lim_{k \to \infty} E(X_{t+k} | X_t, X_{t-1}, ..) = X_t + \left(\sum_{i=1}^{\infty} b_i \right) \varepsilon_t + \left(\sum_{i=2}^{\infty} b_i \right) \varepsilon_{t-1} + ... \tag{2.12}$$

Es pot definir així un nou procés $\widetilde{X}_t = E(X_{t+k} | X_t, X_{t-1}, ..)$ que es coneix com component de llarg termini de X_t. Observi's també que:

$$\widetilde{X}_t - \widetilde{X}_{t-1} = \left(\sum_{i=0}^{\infty} b_i \right) \varepsilon_t \tag{2.13}$$

de manera que $\Delta \widetilde{X}_t$ és un camí aleatori (a no ser que $\sum b_i = 0$). Aquesta expressió recull com es revisa la predicció de la component de llarg termini front a nova informació. Observi's que s'obté el mateix resultat si calculem el límit de l'expressió (2.11):

$$\lim_{k \to \infty} \sum_{i=0}^{k} b_i = \sum_{i=0}^{\infty} b_i = b(1) \tag{2.14}$$

Tal com s'ha comentat abans, alguns autors han proposat usar *b(1)* com a mesura de persistència del procés. Observi's com la component de llarg termini també es pot escriure com:

$$\Delta \widetilde{X}_t = b(1) \varepsilon_t \tag{2.15}$$

La variància de $\Delta \widetilde{X}_t$ que notarem com V_∞ es coneix com variància de llarg termini de la sèrie. És immediat a partir dels resultats anteriors comprovar com aquesta variància és:

$$V_\infty = Var(\Delta \widetilde{X}_t) = b(1)^2 \sigma_\varepsilon^2 \tag{2.16}$$

Aquesta variància a llarg termini pot servir també com a mesura de la persistència de la sèrie. En tot cas, observi's la relació d'aquesta mesura amb *b(1)* analitzada a l'apartat anterior.

Si es compara *b(L) amb b(1)* s'obté:

$$b(L)-b(1)=-(1-L)\left[b_1+b_2(1+L)+b_3(1+L+L^2)+..\right] \tag{2.17}$$

és a dir:

$$b(L)-b(1)=(1-L)b^{*}(L) \tag{2.18}$$

on:

$$b_k^{*}=-\sum_{i=k+1}^{\infty} b_i \tag{2.19}$$

A partir d'aquesta igualtat, l'expressió inicial del procés estocàstic es transforma en:

$$\Delta X_t=b(1)\varepsilon_t+(1-L)b^{*}(L)\varepsilon_t \tag{2.20}$$

de manera que:

$$X_t=b(1)\sum_{i=1}^{t}\varepsilon_i+b^{*}(L)\varepsilon_t+X_0-b^{*}(L)\varepsilon_0 \tag{2.21}$$

Tot i que és habitual suposar que el procés X_t comença en $t=-\infty$, si es pren com a punt de partida el valor $t=0$, es pot canviar la definició de component a llarg termini per:

$$\widetilde{X}_t=b(1)\sum_{i=1}^{t}\varepsilon_i \tag{2.22}$$

D'aquesta manera, l'expressió anterior permet descomposar el procés estocàstic en component a llarg termini i component a curt termini, $b^{*}(L)\varepsilon_t$ i el terme inicial $X_0-b^{*}(L)\varepsilon_0$. Com el terme inicial és independent de la component a llarg termini finalment definida, l'anàlisi d'aquesta component de llarg termini condicional és equivalent a l'anàlisi de la component de llarg termini incondicional. A més, es pot escriure la component de llarg termini com:

$$\widetilde{X}_t=X_t-b^{*}(L)\varepsilon_t=b(1)\sum_{i=1}^{t}\varepsilon_i+X_0-b^{*}(L)\varepsilon_0 \tag{2.23}$$

de manera que segueix complint-se que:

$$(1-L)\widetilde{X}_t=b(1)\varepsilon_t \tag{2.24}$$

Pel que fa a la variància de llarg termini V_∞ hi ha diferents representacions alternatives que en capítols posteriors cal tenir en compte[3]:

1) Recordi's que la densitat espectral de ΔX_t és:

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}\left|\sum_{k=0}^{\infty} b_k e^{-ik\omega}\right|^2 \tag{2.25}$$

de manera que:

$$V_\infty = Var(\Delta\widetilde{X}_t) = 2\pi f(0) \tag{2.26}$$

2) A partir de la definició de la densitat espectral es pot expressar aquesta variància de llarg termini en funció de les autocovariàncies:

$$V_\infty = \gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \tag{2.27}$$

2) Una altra expressió interessant que es pot demostrar a partir de la definició de densitat espectral:

$$Var(\Delta\widetilde{X}_t) = \gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \tag{2.28}$$

3) A més, es pot comprovar com es compleix la següent propietat:

$$\lim_{t\to\infty}\frac{E(X_t^2)}{t} = V_\infty \tag{2.29}$$

En resum, a partir de la funció de resposta a l'impuls, es poden usar com a mesures de persistència les següents:

- el valor de la densitat espectral en la freqüència zero *f(0)*

[3] Per una demostració d'aquestes igualtats es pot consultar Cochrane (1988).

- la resposta del predictor a llarg termini front a una innovació unitària en el moment t que hem notat com $b(1)$.

- $b(1)^2\sigma_\varepsilon^2$, que és la densitat espectral de $(1-L)X_t$ a la freqüència zero i la variància de la innovació de la component camí aleatori:

$$b(1)^2\sigma_\varepsilon^2 = (1+2\sum_{j=1}^{\infty}\rho_j)\sigma_{\Delta X_t}^2 \tag{2.30}$$

Sobre aquestes mesures cal fer alguns comentaris importants:

- Una crítica a aquest tipus d'aproximació és la variabilitat dels resultats segons el model ARMA seleccionat per modelitzar la sèrie diferenciada. Un bon exemple n'és la Taula IV de Campbell i Mankiw (1987) en que l'estadístic b_{80} varia entre 0.004 i 1.812 segons el model usat, per això sembla més interessant alguna aproximació no paramètrica.

- A més, per desenvolupar aquest tipus d'anàlisi no cal especificar un model temporal, es poden usar directament les autocovariàncies estimades i algun algorisme de predicció òptima basat únicament en aquestes autocovariàncies. L'autor d'aquesta tesi no té constància de cap aplicació d'aquest tipus.

2.4. La funció variància-temps

Una altra mesura molt usada per analitzar el grau de persistència en dades temporals és el que es coneix amb el nom de funció variància-temps. Aquesta funció consisteix a calcular per cada retard k, la variància de la diferència entre les dades separades per k unitats de temps, és a dir:

$$\sigma_k^2 = Var(X_t - X_{t-k}) \tag{2.31}$$

Autors com Cochrane (1988), Fama i French (1988) i Huizinga (1986) proposen analitzar la persistència de les sèries macroeconòmiques amb aquesta funció.

Una de les motivacions per usar aquesta funció és que presenta un comportament ben diferent quan les dades observades X_t segueixen un camí aleatori o quan es comporten com un soroll blanc. En el primer cas, es té que $\sigma_k^2=k\sigma_1^2$ de manera que la funció variància-temps creix de manera lineal, en canvi, sota la hipòtesi de soroll blanc, la funció variància-temps és constant e igual a $2\sigma^2$.

De totes maneres, hi ha antecedents molt anteriors en que ja s'usa aquesta funció per mesurar la dependència de llarg termini com poden ser els treballs de Working (1949), Osborne (1959), Sprenkle (1961), Cootner (1964) i Poole (1967). També hi ha avenços posteriors a Young (1971), Mandelbrot (1972), Lo i MacKinlay (1987) i Poterba i Summers (1987) entre altres.

Per estimar aquestes variàncies, sota el supòsit que l'esperança és zero, es pot usar el següent estimador:

$$\widetilde{\sigma}_k^2 = \frac{1}{(T-k+1)}\sum_{t=k}^{T}(X_t - X_{t-k})^2 \tag{2.32}$$

o bé, si cal estimar l'esperança:

$$\widetilde{\sigma}_k^2 = \frac{1}{(T-k+1)}\sum_{t=k}^{T}(X_t - X_{t-k} - k\widetilde{\mu})^2 \tag{2.33}$$

on:

$$\widetilde{\mu} = \frac{1}{T-1}\sum_{t=2}^{T}\left(X_t - X_{t-1}\right) \tag{2.34}$$

Per contrastar la hipòtesi de camí aleatori sembla natural usar com estadístic:

$$R(k) = \frac{k\widetilde{\sigma}_1^2}{\widetilde{\sigma}_k^2} \tag{2.35}$$

Si el procés és un camí aleatori, llavors per cada k:

$$\sigma_k^2 = k\sigma_1^2 \tag{2.35}$$

de manera que *R(k)* és constant. Precisament, aquesta és la hipòtesi que es desitja contrastar. Si es suposa (de manera incorrecta) que $\Delta_1 X_t$ i $\Delta_k X_t$ són independents, llavors es pot usar un test estandar amb distribució *F* de Fisher-Snedecor perquè:

$$\frac{(T-1)\tilde{\sigma}_1^2}{\sigma_1^2} \sim \chi^2_{T-1} \tag{2.36}$$

i:

$$\frac{(T-k)\tilde{\sigma}_k^2}{k\sigma_1^2} \sim \chi^2_{T-k} \tag{2.37}$$

Així, sota el supòsit d'independència l'estadístic proposat seguiria una distribució *F* de Fisher-Snedecor:

$$R(k) = \frac{k\tilde{\sigma}_1^2}{\tilde{\sigma}_k^2} \sim F_{T-1,T-k} \tag{2.38}$$

De totes maneres, el supòsit d'independència no és cert perquè:

$$\Delta_k X_t = \Delta_1 X_t + \Delta_1 X_{t+1} + \ldots + \Delta_1 X_{t+k-1} = \sum_{j=1}^{k} \Delta_1 X_{t+k-j} \tag{2.39}$$

Per tant, cal disposar de valors crítics per poder contrastar la hipòtesi del camí aleatori. Diebold (1989) tabula aquests valors crítics per la seva utilització per contrastar la hipòtesi nul·la de camí aleatori.

Un aspecte destacat de la funció variància-temps és que està molt relacionada amb la variància a llarg termini presentada a l'apartat anterior. Concretament, i seguint la notació usada abans, les diferències d'ordre *k* de la variable observada es poden expressar com:

$$\Delta_k X_t = X_{t+k} - X_t = \sum_{s=1}^{k} \Delta X_{t+s} = k\mu + b(1)\sum_{s=1}^{k} \varepsilon_{t+s} + b^*(L)(\varepsilon_{t+k} - \varepsilon_t) \tag{2.40}$$

de manera que:

$$E(X_{t+k} - X_t - k\mu)^2 = kb(1)^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2\left(\sum_{j=0}^{k-1} b_j^{*2} + 2b(1)\sum_{j=0}^{k-1} b_j^* + \sum_{j=0}^{\infty}\left(b_{j+k}^* - b_j^*\right)^2\right) \tag{2.41}$$

Consideri's ara la variància de les diferències d'ordre k del procés X_t normalitzada per k:

$$V_k = \frac{1}{k} Var(X_{t+k} - X_t) = \frac{1}{k} E(X_{t+k} - X_t - k\mu)^2 \tag{2.42}$$

com el segon terme de l'expressió anterior està acotat a mesura que k tendeix a infinit, s'obté que:

$$\lim_{k \to \infty} V_k = b(1)^2 \sigma_\varepsilon^2 \tag{2.43}$$

Una expressió alternativa i clarificadora per interpretar aquests resultats és la següent[4]:

$$V_k = \frac{1}{k} E\big((\Delta X_{t+1} - \mu) + \ldots + (\Delta X_{t+k} - \mu)\big)^2 = \gamma_0 + 2\sum_{j=1}^{k-1} \frac{k-j}{k} \gamma_j \tag{2.44}$$

on γ_j són les autocovariàncies de ΔX_t. De fet, el límit d'aquesta variància és:

$$\lim_{k \to \infty} V_k = \gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j = 2\pi f(0) = V_\infty \tag{2.45}$$

que coincideix amb la variància a llarg termini analitzada a l'apartat anterior.

Pel resultat anterior es pot estimar V_k i V_∞ a partir d'estimacions de les autocovariàncies. Huizinga (1987) i Campbell i Mankiw (1988) ho fan així. De fet, l'expressió anterior és la fórmula de Bartlett[5] per estimar l'espectre a la freqüència zero. Com les propietats de l'estimador de Bartlett són ben conegudes, això permet conèixer les propietats asimptòtiques de l'estimador de la variància de les k-diferències.

Cochrane (1988) argumenta que en mostra finita és un estimador que té un biaix important i proposa una correcció de l'estimador que és no esbiaixat si el procés estocàstic és un camí aleatori:

$$\widetilde{V}_k = \frac{T-1}{k(T-k)(T-k-1)} \sum_{j=k+1}^{T} \left[X_t - X_{t-k} - \frac{k}{T-1}(X_T - X_1) \right]^2 \tag{2.46}$$

[4] Demostrada a Cochrane (1988).

En efecte, si $X_t = \mu + X_{t-1} + \varepsilon_t$, llavors $E(\widetilde{V}_k) = V_k$, mentre que si $X_t = \mu + X_{t-1} + (1+\theta L)\varepsilon_t$, llavors:

$$E(\widetilde{V}_k) = (1+\theta)^2 \sigma_\varepsilon^2 - \frac{2\theta}{k} \frac{1 + (k^2 \ (T-1)^2) - \left[2k \ (T-1)(T-k-2)\right]}{1 - k \ (T-1)} \sigma_\varepsilon^2 \tag{2.47}$$

Consideri's ara el següent quocient que s'introdueix per a facilitar l'anàlisi de la distribució de l'estadístic resultant:

$$r_k = \frac{V_k}{V_1} \tag{2.48}$$

Com:

$$E(\Delta X_{t+1} - \mu)^2 = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j^2 \tag{2.49}$$

s'obté que el límit de r_k és:

$$r = \lim_{k \to \infty} r_k = \frac{\sum_{j=0}^{\infty} b_j}{\sum_{j=0}^{\infty} b_j^2} = \frac{b(1)}{\sum_{j=0}^{\infty} b_j^2} \tag{2.50}$$

Per tant, si el procés X_t és TS, llavors $b(1)=0$ de manera que $r=0$. En canvi, si és DS, com $b(1)>0$, $r>0$ ja que b_j^2 és convergent. En el cas particular en que el procés és un camí aleatori, s'obté que $b(1)=1$ i que $r=1$. Així, si $r<1$ es diu que la sèrie presenta reversió a la mitjana.

De fet, tal i com es podrà comprovar al llarg de la tesi, el cas TS és un cas extrem de reversió a la mitjana, ja que $r=0$ però TS i reversió a la mitjana no són equivalents.

[5] Vegi's Bartlett (1950).

2.5. El rang reescalat

El primer estadístic usat per mesurar la presència del concepte de memòria llarga és original de Hurst (1951) i molt utilitzat en els treballs de Mandelbrot (1972, 1975). Es tracta de l'estadístic conegut amb el nom de rang reescalat o estadístic R/S. Per introduir aquest estadístic és adequat introduir algunes notacions. Si X_t són les observacions d'un procés estocàstic, es defineix un nou procés X_t^* com la suma acumulada d'aquestes observacions, és a dir:

$$X_t^* = \sum_{j=1}^{t} X_j \tag{2.51}$$

Suposi's també que l'esperança de X_t és coneguda i nul·la. Si es traça una línia des de X_1^* fins X_T^* s'obté una aproximació a la tendència de X_t^*. Una aproximació a l'amplitud de les oscil·lacions de la variable X_t ve definida per la diferència entre el valor màxim i el valor mínim d'aquesta nova sèrie, és a dir:

$$R_T^{(1)} = \max_{0 \leq t \leq T} X_t^* - \min_{0 \leq t \leq T} X_t^* \tag{2.52}$$

Però el més habitual és no conèixer l'esperança de manera que cal ajustar l'estadístic amb estimacions d'aquesta esperança. La manera natural de fer-ho és usar en aquest cas:

$$R_T^{(2)} = \max_{0 \leq t \leq T} \left\{ X_t^* - T\overline{X} \right\} - \min_{0 \leq t \leq T} \left\{ X_t^* - T\overline{X} \right\} \tag{2.53}$$

Evidentment, l'amplitud de les oscil·lacions mesurada pel rang serà major quan major sigui la variància de la sèrie. Així, per estandaritzar aquest estadístic és natural posar-lo en relació a una estimació de la desviació estàndard mostral:

$$S_T = \left\{ T^{-1} \sum_{t=1}^{T} \left(X_t - \overline{X} \right)^2 \right\}^{1/2} \tag{2.54}$$

La idea intuïtiva que motiva la utilització d'aquest estadístic per part dels primers interessats en la mesura de la memòria llarga és la següent. A Hurst (1951) s'associa la noció de memòria llarga amb la presència d'unes autocorrelacions amb una estructura hiperbòlica que compleixi $\gamma_k \sim Ak^{2(H-1)}$ on s'imposa al coeficient H la condició de que $H \in (1/2, 1)$. Doncs bé, si una sèrie compleix aquesta característica és fàcil comprovar que el quocient entre els rang i la desviació estàndard es comporta de manera aproximada com:

$$\frac{R_T}{S_T} \sim CT^H \tag{2.55}$$

on C és una certa constant. Per tant, aquest autor proposa calcular el quocient (2.55) anomenat rang reescalat i a partir d'aquí intentar esbrinar el valor d'H, conegut com *coeficient de Hurst.* Si s'obté un valor estadísticament superior a 1/2, llavors hi ha evidència de memòria llarga.

Hurst proposa la utilització de l'estadístic R/S en la següent aproximació:

$$\log\left[E\left(R_T/S_T\right)\right] \sim constant + H[\log(T)] \tag{2.56}$$

de manera que pot estimar-se el coeficient de Hurst H com:

$$\tilde{H} = \frac{\log\left(R_T \cdot S_T\right)}{\log(T)} \tag{2.57}$$

De totes maneres, altres autors han proposat definir l'estadístic R/S a partir de diferents submostres de les dades per disposar de més informació. Així, si la mostra utilitzada comença en l'observació $t+1$-èssima i té un tamany s, es defineix:

$$R^{(3)}(t,s) = \max_{0\le u\le s}\left\{X^{*}_{t+u} - X^{*}_{t} - u\overline{X}\right\} - \min_{0\le u\le s}\left\{X^{*}_{t+u} - X^{*}_{t} - u\overline{X}\right\} \tag{2.58}$$

Ara bé, com suggereix Mandelbrot (1972) és més recomanable usar en cada moment estimadors de la mitjana a partir de les observacions de la submostra considerada, és a dir:

$$R^{(4)}(t,s) = \max_{0\le u\le s}\left\{X^{*}_{t+u} - X^{*}_{t} - \frac{u}{s}\left(X^{*}_{t+s} - X^{*}_{t}\right)\right\} - \min_{0\le u\le s}\left\{X^{*}_{t+u} - X^{*}_{t} - \frac{u}{s}\left(X^{*}_{t+s} - X^{*}_{t}\right)\right\} \tag{2.59}$$

De la mateixa manera, és recomanable ponderar el rang en cada submostra per una desviació estandar estimada a partir de les dades de la submostra:

$$S(t,s) = \left\{s^{-1}\sum_{u=1}^{s}\left[X_{t+u} - s^{-1}\left(\sum_{u=1}^{s}X_{t+u}\right)\right]^2\right\}^{1/2} = \left\{s^{-1}\sum_{u=1}^{s}X^2_{t+u} - s^{-2}\left(\sum_{u=1}^{s}X_{t+u}\right)^2\right\}^{1/2} \tag{2.60}$$

En aquest cas, enlloc d'usar un quocient i com es disposa de diferents valors de R i S, es pot estimar H com la pendent de realitzar una regressió entre *log(R(t,s)/S(t,s))* i *log(s)*.

2.6. La variància de les sumes parcials

Un altre instrument molt usat en aquest context, també relacionat amb els anteriors, és la variància de les sumes parcials de la sèrie X_t analitzada:

$$S_T = X_1 + X_2 + \ldots + X_T \tag{2.61}$$

La variància d'aquestes sumes $V_T = Var(S_T)$ està molt relacionada amb el grau de memòria de la sèrie X_t.

Sota el supòsit d'estacionarietat es pot demostrar que:

$$V_T = T\sigma_X^2 \left[1 + 2\sum_{k=1}^{T-1} \frac{T-k}{T} \rho_k \right] \tag{2.62}$$

Cochrane (1988) usa aquesta funció com a mesura de persistència. Observi's que asimptòticament es compleix que:

$$\frac{V_T}{T\sigma_X^2} \underset{T\to\infty}{\to} 2\pi f(0) \tag{2.63}$$

límit que coincideix amb la variància a llarg termini V_∞ presentada a l'apartat anterior.

Aquesta mesura de persistència, però, ha estat molt criticada per Quah (1992) i altres autors.

2.7. Consideracions finals

Tal com s'ha comprovat al llarg dels apartats anteriors, no és possible establir una diferenciació clara entre les diferents mesures que s'han utilitzat a la literatura economètrica per mesurar el grau de *persistència* de diferents variables econòmiques, ja que estan molt relacionades les unes amb les altres.

Com es comentava a la introducció d'aquest capítol, l'objectiu dels apartats anteriors era el de mostrar com la noció de *persistència* és una noció que apareix de manera recurrent en diferents àmbits de l'anàlisi econòmica. Així, diferents autors i en diferents moments han utilitzat diferents eines per mesurar un concepte sobre el que no hi ha definicions acceptades de manera universal.

Per tant, previ a l'anàlisi de les conseqüències que té la *memòria* per l'Econometria, és imprescindible establir de manera precisa una classificació de les sèries temporals segons la seva *memòria* que sigui adequada per aquest objectiu, tasca abordada al capítol següent.

Capítol 3

DEFINICIÓ DE LA MEMÒRIA D'UNA SÈRIE TEMPORAL

3.1. Introducció

Els diferents models de sèries temporals s'acostumen a valorar com adequats o no per a les observacions d'una determinada variable en funció de si els residus són soroll blanc. Implícit en aquest plantejament hi ha la concepció de que unes determinades dades no presenten memòria quan són soroll blanc. Aquest és el supòsit ideal en tot tipus d'anàlisi estadística i això fa que aquelles dades que presentin certa memòria necessitin, sovint, d'algun tipus de modelització.

De fet, és probable que hi hagi unanimitat entre els diferents autors sobre aquesta caracterització de l'absència de memòria. El problema sorgeix a l'hora de subdividir en categories les sèries en que es rebutja aquesta incorrelació o absència de memòria.

Sovint, el tipus de memòria s'ha considerat més com a restricció per a la realització d'inferència que com a característica no paramètrica de les dades temporals. En aquest sentit, una de les motivacions més importants d'aquesta classificació és discriminar entre aquelles dades que es poden modelitzar a través de models ARMA, i aquelles que no.

Així, es podria dir que la primera i més primària classificació de les sèries temporals en funció de la seva memòria, és la distinció entre les sèries que són estacionàries i ergòdiques i per tant, que permeten aplicar-hi les tècniques habituals basades en les correlacions (per exemple pels economistes) o en el periodograma (sobretot pels enginyers) i les sèries en que no són aplicables directament aquestes tècniques. Segons aquesta concepció hi hauria memòria llarga quan les dades presenten tendències, estacionalitat, no estacionarietat o components deterministes.

Però aquesta és una classificació molt vaga i que en tot cas està causada per la impossibilitat de modelitzar aquests comportaments i no per raons objectives.

En tot cas, s'ha de destacar la presència d'autors com Mandelbrot que són molt crítics amb aquesta perspectiva per a l'anàlisi de dades econòmiques temporals ja que creuen que el correlograma i el periodograma no són eines adequades per a l'anàlisi d'aquest tipus de sèries.

Històricament, es produeix un salt qualitatiu en la concepció del grau de memòria a partir de la proposta de diferenciar aquelles dades que semblen no estacionàries. Aquesta classificació porta a distingir un tipus de memòria associada als models ARMA d'una altre associada amb els models ARIMA. De totes maneres, tal com s'ha argumentat prèviament, entre les característiques d'un model ARMA i les d'un model ARIMA hi ha encara massa possibilitats com per a que aquesta sigui una bona classificació.

En tot cas, aquestes no són classificacions formals i a continuació es centra l'anàlisi a algunes possibilitats de precisar més aquesta distinció. Així, diferents autors han proposat diferents aproximacions al concepte de memòria. A continuació es presenten cada una d'aquestes aproximacions per relacionar-les entre sí i, finalment, escollir la que en aquesta tesi es pren com a definició de referència per a l'anàlisi de les conseqüències que té per a l'anàlisi economètrica el tipus de memòria. També cal destacar que la intenció és la de treballar amb conceptes el màxim de generals possibles.

3.2. Tipus de memòria a partir de l'estructura de correlacions

El concepte de memòria està indefugiblement associat amb la importància del grau d'associació entre observacions força distants en el temps. Però la mesura d'aquesta associació no pot basar-se directament a partir dels coeficients de correlació mostral ja que es tracta d'un estadístic amb propietats que depenen de les pròpies característiques d'aquesta dependència. Per tant, cal fer aquesta anàlisi amb cura.

Per valorar les definicions que es presenten a continuació, cal situar-les amb una certa perspectiva històrica. En aquest sentit, una de les bases del modern anàlisi de sèries temporals es troba en el treball de Schuster (1898) que és un dels primers en considerar les observacions d'una sèrie temporal X_t com una part finita d'un procés estocàstic infinit. En aquest context es proposa analitzar aquest tipus de dades a partir del següent estadístic:

$$f_T(\omega) = \frac{1}{T}\left|\sum_{t=1}^{T} X_t e^{-it\omega}\right| \qquad (3.1)$$

Aquest estadístic és, encara avui, una eina fonamental per a l'anàlisi de sèries temporals[1]. La variable ω s'interpreta com a freqüència i s'acostuma a representar gràficament l'estadístic $f_T(\omega)$ per diferents valors d'aquesta freqüència.

Aquest gràfic és molt erràtic, de manera que el concepte de densitat espectral sorgeix de manera intuïtiva prenent el límit de la funció (3.1) quan el número d'observacions T augmenta. Es defineix així la funció $f(\omega)$:

$$f(\omega) = \lim_{T\to\infty} f_T(\omega) \qquad (3.2)$$

De totes maneres, la definició (3.2) és sols vàlida de manera intuïtiva i per tant, cal definir amb certa precisió aquest límit per a garantir la seva existència.

Wiener (1930) proposa resoldre aquest problema a través de la definició de la funció d'autocovariàncies mostral. A partir d'aquell moment, aquesta funció guanya gran popularitat com a eina fonamental per a l'anàlisi de tota sèrie temporal. Si es disposa d'un conjunt de T observacions X_t es defineix cada una de les autocovariàncies mostrals com:

$$C_T(h) = \frac{1}{T}\sum_{t=1}^{T-h}(X_{t+h} - \overline{X})(X_t - \overline{X}) \qquad (3.3)$$

per diferents valors del retard h. L'argument de Wiener es basa en la demostració de que aquesta funció $C_T(h)$ és la transformada de Fourier de $f_T(\omega)$:

[1] En la terminologia actual s'utilitza el periodograma $I(\omega)$, que no és més que la funció $f_T(\omega)$ normalitzada, de manera que $I(\omega)=f_T(\omega)/2\pi$.

$$C_T(h) = \int_{-\pi}^{\pi} f_T(\omega) e^{ih\omega} d\omega \qquad (3.4)$$

Així, sota el supòsit de que existeix el límit d'aquestes autocovariàncies:

$$C(h) = \lim_{T \to \infty} C_T(h) \qquad (3.5)$$

coneguda amb el nom de funció d'autocovariàncies asimptòtica, queda garantida l'existència del límit definit a (3.2).

Per obviar els problemes derivats de la dependència d'aquests estadístics front diferents graus de dispersió de les dades es defineix la funció de correlacions asimptòtica com:

$$r(h) = \frac{C(h)}{C(0)} \qquad (3.6)$$

Aquesta funció és també el límit de les funcions d'autocorrelació mostral:

$$r_T(h) = \frac{C_T(h)}{C_T(0)} \qquad (3.7)$$

El gràfic d'aquest instrument conegut amb el nom de correlograma és molt usat com a mètode d'identificació del comportament de les dades.

Un model de probabilitat senzill sota el qual existeix la funció d'autocovariàncies asimptòtica és el d'un procés estocàstic gaussià amb esperança zero i estacionarietat de segon ordre amb funció d'autocovariàncies $\gamma_h = E[X_{t+h} X_t]$. Sota aquest model es compleix que $\gamma_h = C(h)$.

De totes maneres, i tot i que sembli sorprenent, no n'hi ha prou amb el supòsit d'estacionarietat per garantir que la funció d'autocovariàncies mostral tendeixi a la funció d'autocovariàncies γ_h.

Quan es compleix que les autocovariàncies mostrals tendeixen en probabilitat[2] a γ_h:

$$C_T(h) \xrightarrow{p} \gamma_h \tag{3.8}$$

i per tant es tracta d'estimadors consistents d'aquesta funció, es diu que la sèrie temporal és ergòdica.

Es pot comprovar com una condició necessària i suficient d'ergodicitat en el cas de sèries gaussianes i estacionàries és que:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{h=1}^{T} \rho^2(h) = 0 \tag{3.9}$$

Tenint en compte això sembla natural la següent classificació per als processos estacionaris:

Classificació[3] A

*Definició A1**. Es diu que una sèrie temporal estacionària no presenta memòria si[4]:

$$\frac{1}{T}\sum_{h=1}^{T} \rho^2(h) = 0 \tag{3.10}$$

*Definició A2**. Es diu que una sèrie temporal estacionària presenta memòria curta si:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{h=1}^{T} \rho^2(h) = 0 \tag{3.11}$$

*Definició A3**. Es diu que una sèrie temporal estacionària presenta memòria llarga si:

---

[2] Si $\{b_T(\omega)\}$ és una successió de variables aleatories, es diu que $b_T(\omega)$ convergeix en probabilitat si existeix un valor b tal que per a qualsevol $\varepsilon>0$, es compleix que $P\left[|b_T(\omega)-b|<\varepsilon\right] \underset{T\to\infty}{\longrightarrow} 1$. Si aquest és el cas, s'utilitza la següent notació: $b_T(\omega) \xrightarrow{p} b$.

[3] Les definicions assenyalades amb un asterisc (*) són provisionals fins donar les definicions definitives al final del capítol.

[4] Recordi's que si una sèrie és estacionària (de segon ordre), es compleix sempre que $\sum_{h=1}^{T} \rho^2(h) < \infty$.

$$\lim_{T \to \infty} \frac{1}{T} \sum_{h=1}^{T} \rho^2(h) \neq 0 \qquad (3.12)$$

Noti's que la propietat (3.10) implica (3.11) de manera que les definicions A0 i A1 no són excloents sinó que la no presència de memòria és un cas particular de memòria curta. Aquesta característica és comuna a altres definicions que s'analitzen a continuació, de manera que molts autors en centren en precisar la distinció entre memòria curta i memòria llarga sense presentar definicions explícites de l'absència de memòria.

En tot cas, probablement trobaríem unanimitat en que són equivalents l'absència de memòria i comportament soroll blanc. Un cop fixada aquesta caracterització també són equivalents la incorrelació de les observacions i la propietat (3.10).

Cal insistir en que la definició anterior de memòria llarga està associada estrictament al concepte d'ergodicitat. És per això, que aquesta definició presenta varis aspectes negatius:

- D'una banda, es tracta d'una classificació inaplicable a la pràctica, ja que és impossible contrastar la condició (3.11) amb una sola realització.

- A més, les conseqüències de la no ergodicitat són tan greus que invaliden la majoria d'eines habituals a l'anàlisi de dades econòmiques temporals.

- Com s'ha mostrat al capítol 1, la condició de sumabilitat absoluta no és equivalent a l'ergodicitat, de manera que hi ha la possibilitat que una sèrie sigui ergòdica però amb major memòria de la que es pot modelitzar a través d'un model ARMA, possibilitat que la Classificació A no considera.

Sovint, però, es dona una altra definició més operativa en termes de la sumabilitat absoluta de les autocorrelacions ja que aquesta és una condició suficient d'ergodicitat de segon ordre.

Així, seguint McLeod i Hipel (1978) es pot establir la següent classificació:

Classificació B

Definició B1*. Es diu que una sèrie temporal estacionària presenta memòria curta si:

$$\sum_{h=1}^{\infty} |\rho_k| < \infty \tag{3.13}$$

Definició B2*. Es diu que una sèrie temporal estacionària presenta memòria llarga si:

$$\sum_{h=1}^{\infty} |\rho_k| = \infty \tag{3.14}$$

Noti's com la condició (3.11) implica la condició (3.13) de manera que una sèrie temporal amb memòria curta segons la definició A també ho és segons la definició B tot i que la implicació contrària no és certa, de manera que la classificació B considera com memòria llarga sèries que la classificació A considera com memòria curta.

Una altra possibilitat consisteix a definir el concepte de memòria llarga directament en termes de la funció d'autocovariàncies, limitant per tant la classificació a dades estacionàries. Així, per exemple, a Helson i Sarason (1967) es donen les següents definicions:

Classificació[5] C

Definició C1*. Es diu que una sèrie temporal estacionària presenta memòria llarga si:

$$\gamma_h \sim g(h)h^{\lambda} \tag{3.15}$$

quan $k \to \infty$, $\lambda > -2$ i $g(\cdot)$ és una funció amb variació suau a l'infinit.

Es diu que una funció $g(\cdot)$ és de variació regular a l'infinit amb índex α si és positiva, mesurable i compleix la següent propietat:

[5] A Fox i Taqqu (1986) i Robinson (1995) es donen definicions similars basades en el comportament de les autocorrelacions. En aquests tipus de definicions és habitual usar els conceptes *long range dependence* o *strong dependence* enlloc de *long memory*.

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = x^{\alpha} \tag{3.16}$$

per qualsevol $x>0$. Una funció és de variació suau a l'infinit si és de variació regular amb $\alpha=0$, és a dir, que asimptòticament $g(\cdot)$ es comporta com una constant. Un exemple molt senzill de funció de variació suau a l'infinit és $g(x)=log(x)$.

*Definició C2**. Si no és així, es diu que la sèrie temporal presenta memòria curta.

3.3. Tipus de memòria a partir del comportament de les sumes parcials

Evidentment, un dels problemes de les definicions de memòria llarga presentades fins al moment és la necessitat de suposar estacionarietat de segon ordre. Així com recentment s'ha popularitzat en econometria la teoria asimptòtica basada en les condicions de mixtura[6] per imposar restriccions de memòria a les dades enlloc d'imposar la condició d'ergodicitat, hi ha també la possibilitat d'usar aquest concepte per formalitzar la idea de memòria llarga.

Així, Rossenblat (1956) usa el concepte de mixtura forta per definir el concepte de *long range dependence*. Per fer-ho utilitza el comportament de les sumes parcials d'una sèrie temporal. La idea intuïtiva que hi ha a darrera d'aquesta definició és ben senzilla. Si les observacions fossin independents, al afegir una nova observació s'afegeix a la variància total la variància corresponent a la nova observació. Així, si a més d'independents tenen totes les observacions igual variància, la variància total creix de manera lineal a mesura que augmenta el nombre d'observacions, és a dir, la variància total normalitzada pel nombre d'observacions és constant. En canvi, si la variància total normalitzada no està acotada, això vol dir que la dependència entre les observacions és ilimitada.

[6] Una excel·lent recopilació d'aquesta teoria asimptòtica es pot consultar a White (1984).

Més formalment, si es disposa de T observacions d'una sèrie temporal X_t, es defineixen les sumes parcials d'aquesta sèrie com:

$$S_T = \sum_{t=1}^{T} X_t \tag{3.17}$$

Noti's com el comportament asimptòtic d'aquestes sumes parcials aporta informació sobre les propietats de l'estadístic *mitjana mostral* ja que $\overline{X}_T = S_T / T$. Concretament, si el procés X_t és estacionari, la variància d'aquestes sumes parcials sempre existeix i ve donada per:

$$Var(X_T) = T \sum_{j=1-T}^{T-1} \left(1 - \frac{|j|}{T}\right) \gamma_j \tag{3.18}$$

Sota molts tipus de dependència dèbil de les dades, per exemple model ARMA, es demostra el següent Teorema Central del Límit:

$$T^{1/2} S_T \Rightarrow N(0, 2\pi f(0)) \tag{3.19}$$

Segons Rossenblat (1956) si la sèrie X_t compleix la condició de mixtura forta, es diu que la sèrie presenta *short range dependence*. Sota la condició de mixtura forta, es pot definir el concepte de variància de llarg termini de la sèrie X_t com:

$$\sigma^2 = \lim_{T \to \infty} E\left[\frac{S_T^2}{T}\right] \tag{3.20}$$

Aquest resultat suggereix la següent classificació:

Classificació D.

Definició D1*. Una sèrie temporal X_t presenta memòria curta si existeix $\lim_{T \to \infty} E(T^{-1} S_T^2)$ i aquest límit és diferent de zero.

Definició D2*. Altrament, es diu que la sèrie temporal X_t presenta memòria llarga

A més, també es compleix sota la condició D1 que:

$$\left[1 / \sigma T^{1/2}\right] S_{[rT]} \Rightarrow B(r) \quad \forall \quad r \in [0,1]$$

Aquesta definició està relacionada amb l'estacionarietat, però requereix l'existència de moments d'ordre superior.

Helson i Sarason (1967) demostren que si una sèrie compleix la condició (3.15) amb $\lambda > -2$ i és de memòria llarga segons la classificació C, llavors no compleix la condició de mixtura forta i també és de memòria llarga segons la classificació D. De totes maneres, la implicació inversa no és certa ja que la condició de mixtura forta és més general que la condició (3.15).

3.4. Tipus de memòria a partir de la densitat espectral

Tota sèrie estacionària, sota algunes condicions de regularitat addicionals, té una densitat espectral associada. A més, el periodograma (equivalent mostral de la densitat espectral) també reflexa la forta dependència en moltes de les dades econòmiques a través d'un predomini de les més baixes freqüències[7]. De fet, com la freqüència zero (o les properes) s'associen a comportaments tendencials, això és indicador d'una memòria gran.

Per tant, aquest sembla també un entorn idoni per a la classificació de la memòria. A més, tal com s'ha argumentat abans, hi ha una relació molt estreta entre periodograma i autocovariàncies mostrals.

A la teoria clàssica és habitual el supòsit que la densitat espectral sigui finita en totes les freqüències (una condició suficient per això és la sumabilitat de les autocovariàncies). Però, de manera intuïtiva, per a que una sèrie sigui estacionària en termes espectrals no cal que l'espectre sigui finit, n'hi ha prou amb que l'àrea delimitada per l'espectre sigui finita. És a dir, sempre que:

[7] En sèries amb un fort comportament estacional, també les freqüències estacionals acumulen bona part de la informació que conté la sèrie. En aquest cas, hi ha la possibilitat que la sèrie presenti memòria llarga no sols en la freqüència zero, sino també en alguna d'aquestes freqüències estacionals.

$$\int_{-\pi}^{\pi} f(\omega)d\omega < \infty \tag{3.21}$$

llavors existeix un procés estacionari amb funció d'autocovariàncies definida per:

$$\gamma_k = \int_{-\pi}^{\pi} f(\omega)e^{i\omega k} \tag{3.22}$$

Com a mesura del grau de memòria Parzen (1981) proposa usar el rang de la densitat espectral en logaritmes:

$$\upsilon = \max_{\omega} f(\omega) - \min_{\omega} f(\omega) \tag{3.23}$$

Aquest criteri porta a la següent classificació:

Classificació E.

Definició E1*. Una sèrie temporal X_t estacionària no conté memòria si $\upsilon=0$.

Definició E2*. Una sèrie temporal X_t estacionària es diu que conté memòria curta si $0<\upsilon<\infty$.

Definició E3*. Una sèrie temporal X_t estacionària es diu que conté memòria llarga si $\upsilon=\infty$.

De totes maneres, és més habitual usar un enfocament semiparamètric. Així, si es suposa que el predomini de la variació de la variable es troba al voltant de la freqüència zero, un comportament funcional explícit senzill que permet reproduir un predomini d'aquestes freqüències és la següent representació:

$$f(\omega) = \omega^{-2\delta} L(\omega) \tag{3.24}$$

En aquest cas és diu que $f(\omega)$ és una funció de variació regular[8] a la freqüència 0 i a la representació (3.24) se l'anomena representació regular d' $f(\omega)$. El coeficient δ es coneix amb el nom[9] d'*Índex del grau de memòria* i permet obtenir una nova classificació[10]:

Classificació F.

Definició F1*. Si $\delta=0$ es diu que la sèrie és de memòria curta.

Definició F2*. Si $\delta\neq 0$ es diu que la sèrie és de memòria llarga.

Noti's com en aquesta classificació no es distingeix el cas en que no hi ha memòria. En tot cas, alguns autors fan també una distinció entre el cas $\delta>0$ al que es refereixen com de memòria llarga i el cas $\delta<0$ en que es fa referència a memòria intermèdia[11].

Sovint, es fa el supòsit més restrictiu[12] que al voltant de la freqüència zero es té que[13]:

---

[8] En general, es diu que una funció qualsevol $g(x)$ és de variació regular al voltant d'x_0 si existeix α tal que $g(x)=(x-x_0)^\alpha h(x)$ i:

1) En un entorn a la dreta d'x_0, $(x-x_0)^\alpha h(x)$ és creixent i $(x-x_0)^{-\alpha} h(x)$ és decreixent.
2) En un entorn a l'esquerra d'x_0, $(x_0-x)^\alpha h(x)$ és decreixent i $(x_0-x)^{-\alpha} h(x)$ és creixent.

I llavors es diu que g(x) té índex de variació regular igual a α. En aquest cas, òbviament $\alpha=2\delta$.

[9] Alguns autors com per exemple Parzen (1981) usen una notació diferent, concretament $f(\omega) = \omega^{-2\delta} L(\omega)$ per a que el procés sigui estacionari sempre que $\delta<1$ i llavors s'anomena al paràmetre δ *Índex de variació regular*. A la tesi s'ha optat per usar una notació diferent per a que en models ARFIMA o ARIMA es compleixi que $d=\delta$.

[10] Alguns dels mètodes semiparamètrics d'estimació presentats al capítol 7 es basen en hipòtesis d'aquest tipus.

[11] En aquest cas, la sèrie no presenta forta persistència sinó que les correlacions alternen el seu signe. De fet, alguns autors anomenen aquest cas com d'*antipersistència*.

[12] Vegi's, per exemple Lobato i Robinson (1996).

[13] Formalment, s'usa a la tesi la notació $f(x) \underset{x\to x_0}{\sim} g(x)$ com abreviació de que $\lim_{x\to x_0} \frac{f(x)}{g(x)} = C$ on C és una constant no nul·la.

$$f(\omega) \underset{\omega \to 0^+}{\sim} C\omega^{-2\delta} \tag{3.25}$$

De fet, alguns processos que compleixen aquesta propietat, també satisfan que[14]:

$$\gamma(h) \underset{h \to \infty}{\sim} Dh^{2\delta-1} \tag{3.26}$$

El valor per $\omega=0$ de la densitat espectral pot ser infinit per $\delta=0$. Això passa si $f(\omega) \sim (\log \omega)^2$ quan $\omega \to 0^+$ que es correspon amb un comportament asimptòtic de les correlacions del tipus $\gamma(h) \sim \dfrac{\log h}{h}$.

De totes maneres, les definicions a partir de la densitat espectral i de la funció d'autocovariàncies no són equivalents. La característica més important a tenir en compte és que les basades en les autocovariàncies necessiten del supòsit de que totes aquestes autocovariàncies són positives.

3.5. Tipus de memòria a partir dels errors de predicció

Una de les principals aplicacions de l'anàlisi de sèries temporals és el càlcul de prediccions. Així, sembla també coherent establir una classificació de la persistència o memòria que presenten les dades temporals a partir del comportament d'aquestes prediccions.

La justificació d'aquest enfocament és molt propera conceptualment a la utilització de la funció de resposta a l'impuls per mesurar el grau de memòria. La principal diferència entre ambdós enfocaments es troba en el fet que la funció de resposta a l'impuls és un instrument estadístic associat als models ARMA, mentre que es pot analitzar el comportament de les prediccions òptimes d'una sèrie temporal sense necessitat de restringir l'anàlisi a models ARMA.

[14] De fet, la densitat espectral anterior amb $\delta>0$ és integrable si i només si $0<\delta<1/2$.

Concretament, si es disposa d'observacions d'una variable X_t, i es desitja predir un cert valor d'aquesta variable a partir dels valors passats, el predictor òptim és:

$$\widetilde{X}_t = E\left[X_t \middle| X_{t-1} \ldots\right] \tag{3.27}$$

La successió formada per els errors de predicció: $e_t = X_t - \widetilde{X}_t$ forma una sèrie temporal soroll blanc amb esperança zero i una certa variància. Un cop normalitzada aquesta variància s'obté una mesura del grau de predictibilitat de la sèrie X_t:

$$\sigma_\infty^2 = \frac{E\left[|e_t|^2\right]}{E\left[|X_t|^2\right]} = \frac{E\left[|e_t|^2\right]}{\gamma(0)} \tag{3.28}$$

A la pràctica, no és possible conèixer aquesta variància perquè sols es disposa d'un nombre finit d'observacions. Així, la predicció òptima basada en un nombre finit d'observacions passades és[15]:

$$\widetilde{X}_t^m = a_1 X_{t-1} + \ldots + a_m X_{t-m} \tag{3.29}$$

amb errors de predicció:

$$e_t^m = X_t - \widetilde{X}_t^m \tag{3.30}$$

i una variància dels error normalitzada:

$$\sigma_m^2 = \frac{E\left[|e_t^m|^2\right]}{\gamma(0)} \tag{3.31}$$

Es pot comprovar, com aquestes variàncies tendeixen a la variància del predictor amb infinites observacions. Una alternativa per calcular aquest estadístic és usar la coneguda fórmula de Kolmogorov basada en la densitat espectral:

$$\log \sigma_\infty^2 = \int_0^{2\pi} \log f(\omega) d\omega \tag{3.32}$$

A la pràctica, és habitual estimar la variància σ^2_∞ com:

[15] On els coeficients a_j es poden calcular, per exemple, a través de l'algorisme de Durbin-Levinson. Per una introducció a aquest algorisme vegi's l'Annex 6.1 de la tesi.

$$\tilde{\sigma}^2_{\infty} = \tilde{\sigma}^2_m \tag{3.33}$$

Alguns dels criteris proposats per seleccionar la m adequada com pot ser el AIC d'Akaike (1974, 1977) o el CAT de Parzen es basen en el concepte d'informació que és una de les altres possibilitats d'establir una classificació del grau de memòria que s'analitza a l'apartat següent. En tot cas, aquest estadístic permet establir la següent classificació:

Classificació G.

*Definició G1**. Si $\sigma^2_{\infty}=0$ es diu que la sèrie no té memòria.

*Definició G2**. Si $0<\sigma^2_{\infty}<1$ es diu que la sèrie és de memòria curta.

*Definició G3**. Si $\sigma^2_{\infty}=1$ es diu que la sèrie és de memòria llarga.

3.6. Tipus de memòria a partir del concepte d'informació

La divergència en termes d'informació d'una densitat de probabilitats estimada *g(x)* respecte a la vertadera densitat *f(x)* es defineix com:

$$I(f;g) = \int_{-\infty}^{\infty} \left\{ -\log \frac{g(x)}{f(x)} \right\} f(x)dx \tag{3.34}$$

Aquesta informació té una descomposició molt important en termes d'entropia:

$$I(f;g) = H(f;g) - H(f) \tag{3.35}$$

on *H(f)* és l'entropia de la densitat vertadera:

$$H(f) = H(f;f) = \int_{-\infty}^{\infty} \{-\log f(x)\} f(x)dx \tag{3.36}$$

i H(f;g) és l'entropia creuada:

$$H(f;g) = \int_{-\infty}^{\infty} \{-\log g(x)\} f(x)dx \tag{3.37}$$

La informació $I(Y|X)$ d'una variable aleatòria continua Y en un vector aleatori X es defineix com:

$$I(Y|X) = I(f_{Y|X}; f_Y) \tag{3.38}$$

L'entropia de la variable Y és:

$$H(Y) = H(f_Y) \tag{3.39}$$

i l'entropia d'Y condicionada al vector X:

$$H(Y|X) = H(f_{Y|X}) \tag{3.40}$$

Així, es pot trobar la següent descomposició:

$$I(Y|X) = H(Y) - H(Y|X) \tag{3.41}$$

En el cas que el vector X està format per dos variables aleatòries, es té que:

$$I(Y|X_1; X_1, X_2) = H(f_{Y|X_1}) - H(f_{Y|X_1,X_2}) = H(Y|X_1) - H(Y|X_1, X_2) \tag{3.42}$$

i de l'aplicació d'aquesta igualtat es pot comprovar com:

$$I(Y|X_1; X_1, X_2) = I(Y|X_1, X_2) - I(Y|X_1) \tag{3.43}$$

Quan X i Y són normals multivariants, si σ^2_Y i $\sigma^2_{Y|X}$ són les variàncies de Y i de Y condicionada a X, respectivament:

$$H(Y) = \frac{1}{2}\log \sigma_Y^2 + \frac{1}{2}(1 + \log 2\pi) \tag{3.44}$$

$$H(Y|X) = \frac{1}{2}\log \sigma_{Y|X}^2 + \frac{1}{2}(1 + \log 2\pi) \tag{3.45}$$

i:

$$I(Y|X) = -\frac{1}{2}\log\left(\sigma_Y^{-2}\sigma_{Y|X}^2\right) \tag{3.46}$$

En el context de la teoria de la informació, es pot tractar el problema de la predicció en sèries temporals a través de la definició de l'estadístic següent:

$$I_m = I(X_{m+1}|X_1,\ldots,X_m) \tag{3.47}$$

L'estadístic I_m mesura el grau d'influència de les dades passades sobre un valor concret de la sèrie X_t. Així, es pot usar el seu límit $I_\infty = \lim_{m\to\infty} I_m$ per classificar les sèries temporals en tres categories

Classificació H.

*Definició H1**. Si $I_\infty=0$ es diu que la sèrie no té memòria.

*Definició H2**. Si $0<I_\infty<\infty$ es diu que la sèrie és de memòria curta.

*Definició H3**. Si $I_\infty=\infty$ es diu que la sèrie és de memòria llarga.

Observi's que aquesta classificació és coherent amb l'anterior ja que per a un procés gaussià estacionari es compleix que:

$$I_\infty = -\frac{1}{2}\log\sigma_\infty^2 \tag{3.48}$$

i per tant, quan $\sigma^2_\infty=1$ llavors es compleix que $I_\infty=\infty$ mentre que si $0<\sigma^2_\infty=1$ llavors $0<I_\infty<\infty$.

3.7. Definició de la memòria d'una sèrie temporal

Cada un dels plantejaments esmentats fins al moment en relació al concepte de memòria és parcial i es centra més en l'objectiu concret d'una certa aplicació que en plantejar de manera general una definició i categorització dels diferents tipus de memòria que poden presentar les dades temporals.

Per donar una visió global del concepte de memòria cal retornar al punt de partida de l'anàlisi de sèries temporals, és a dir, la impossibilitat de mantenir el supòsit d'independència entre les observacions d'una mateixa variable mesurada en diferents moments del temps. Sota aquest punt de vista, la tasca principal de l'anàlisi de sèries temporals consisteix a proporcionar tècniques i models que permetin la modelització més o menys explícita d'aquest correlació. Així, una de les eines clàssiques i bàsiques a aquest tipus d'anàlisi és el correlograma o successió de coeficients d'autocorrelació.

Tot i que es tracta d'un instrument que sols té sentit des d'un punt de vista estadístic sota el supòsit d'estacionarietat de segon ordre, s'acostuma a usar sense aquesta restricció, entre altres coses per detectar problemes de no estacionarietat i té l'avantatge principal en la seva fàcil interpretació.

Atès que la tesi està centrada en les sèries temporals econòmiques, cal preguntar-se a continuació quines són les característiques més importants dels correlogrames habituals en aquest tipus de sèries. Doncs bé, la característica més rellevant és el fet que presenten gran nombre de coeficients no nuls i que presenten un decreixement lent a mesura que augmenta el retard temporal considerat. De fet, aquesta és la traducció en el domini temporal de les característiques destacades per Granger (1966) en el domini de les freqüències.

Noti's que la relació intuïtiva entre aquest instrument estadístic i l'objectiu de la tesi és ben clar, ja que en la mesura que aquestes correlacions siguin importants fins i tot per retards elevats, la sèrie presentarà una memòria gran. Atès per tant, que una eina tan senzilla suggereix que la memòria en les dades econòmiques és gran, la passa següent en la nostra argumentació és la recerca d'alguna estructura teòrica del correlograma que pugui semblar-se a aquests correlogrames empírics.

Recordi's en primer lloc que tot correlograma ha de complir dos condicions derivades de les propietats dels moments de segon ordre sota la hipòtesi d'estacionarietat. Així, si X_t és una sèrie estacionària de segon ordre i es nota per $\rho(k)$ la correlació entre les observacions separades per k unitats de temps, es compleix:

1. que $\rho(0)=1$ i,
2. quan $k\rightarrow\infty$ es compleix que $\rho(k)\rightarrow 0$.

Hi ha infinitat de successions matemàtiques que compleixen aquestes condicions i per tant són candidates a modelitzar les autocorrelacions de sèries temporals però entre les més senzilles hi destaquen les següents[16]:

1. Estructura lineal: $\rho(k)=\begin{cases}1-k(1-\alpha) & 0\le k\le 1/(1-\alpha)\\ 0 & k>1/(1-\alpha)\end{cases}$ on $0\le\alpha\le 1$.

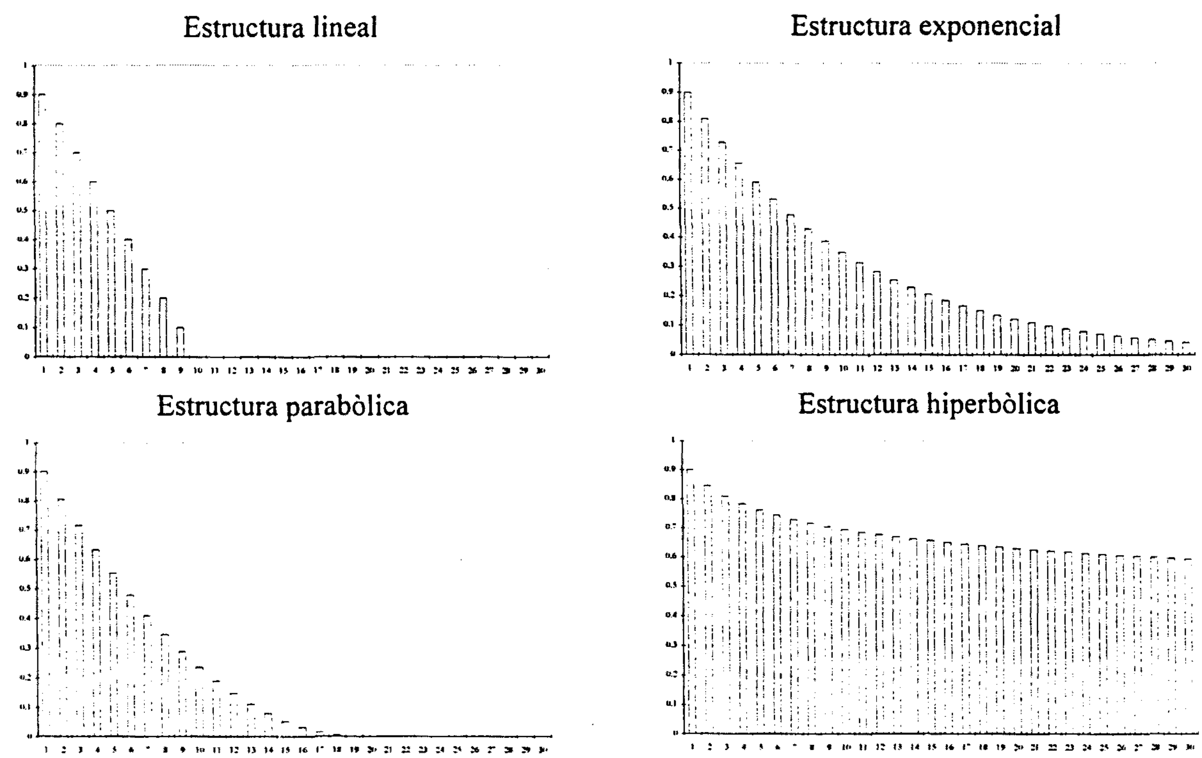
2. Estructura parabòlica: $\rho(k)=\begin{cases}\dfrac{(1-\beta)^2}{4}k^2-(1-\beta)k+1 & 0\le k\le\dfrac{2}{1-\beta}\\ 0 & k>\dfrac{2}{1-\beta}\end{cases}$ on $0\le\beta\le 1$.

3. Estructura exponencial: $\rho(k)=\gamma^k$ on $0\le\gamma\le 1$.

4. Estructura hiperbòlica: $\rho(k)=(1+k)^{\delta-1}$ on $0\le\delta\le 1$.

[16] S'han parametritzat totes les funcions de manera que els paràmetres α, β, γ i $\delta\in[0,1]$, que valors de $\alpha=0$, $\beta=0$, $\gamma=0$ i $\delta=0$ corresponguin amb una absència de correlacions i per tant situació en que les observacions de la serie són independents mentre que els casos $\alpha=1$, $\beta=1$, $\gamma=1$ i $\delta=1$ es corresponguin a un correlograma identicament igual a la unitat, $\rho(k)=1$ per $\forall k$.

Gràfic 3.1 Diferents estructures d'autocorrelació en sèries temporals on $\rho(1)=0.9$.



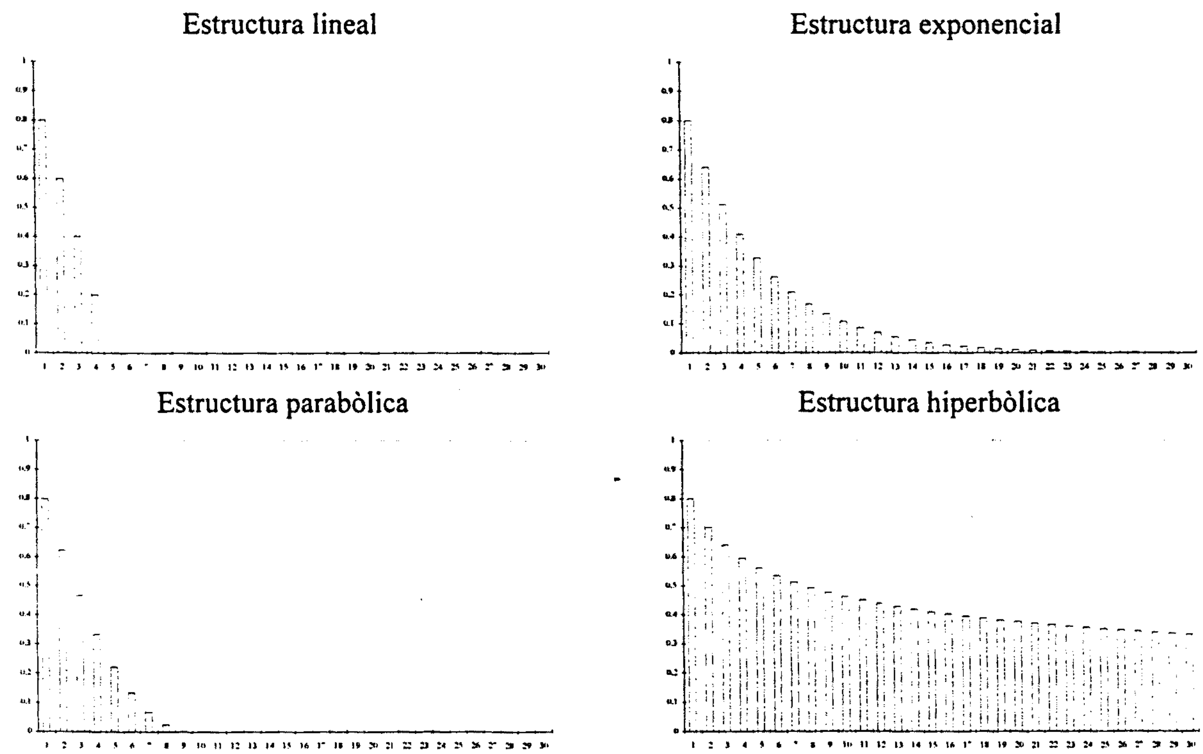
Els gràfics 3.1 i 3.2 mostren una comparació entre aquestes quatre estructures d'autocorrelacions per una correlació de primer ordre fixada per a fer possible aquesta comparació[17].

Observi's com les dos primeres estructrures, la lineal i la parabòlica presenten el problema de que tendeixen ràpidament a zero, i ho fan no de manera asimptòtica sinó de manera exacte, és a dir, a partir d'un cert retard k_0, es compleix $\rho(k)=0$ per $k>k_0$. De fet, aquest és un comportament poc raonable per les dades econòmiques ja que no hi ha raons per pensar (almenys en la majoria de variables econòmiques) que hi hagi alguna dependència entre les

[17] Fixat el valor del primer coeficient d'autocorrelació, $\rho(1)$, es poden calcular els paràmetres corresponents a cada un dels models com $\alpha=\rho(1)$, $\beta = 2\sqrt{\rho(1)} - 1$, $\gamma=\rho(1)$ i $\delta=1+\log_2\rho(1)$.

observacions separades per menys de, per exemple 18 mesos, i a partir d'aquí que aquesta dependència desapareix totalment.

Gràfic 3.2 Diferents estructures d'autocorrelació en sèries temporals on $\rho(1)=0.8$.



En canvi, l'estructura exponencial i hiperbòlica presenten un decreixement cap a zero de manera asimptòtica, característica que sembla més raonable en el cas de variables econòmiques. La diferència fonamental entre l'estructura exponencial i l'estructura hiperbòlica es troba en la velocitat d'aquesta convergència cap al valor zero.

Tot i la fàcil interpretació del correlograma es tracta sols d'una descripció de les correlacions però no d'una modelització explícita del comportament de la sèrie temporal. En aquest sentit, Yule (1927) presenta la idea que les observacions d'una sèrie temporal amb valors altament dependents poden ser conceptualitzats com generats per una successió de shocks ε_t independents i posteriorment un filtre lineal[18].

$$\varepsilon_t \longrightarrow \boxed{\text{FILTRE LINEAL}} \longrightarrow X_t$$

Aquesta idea queda més tard formalitzada a través del Teorema de representació de Wold (1943) que suposa una aportació fonamental ja que demostra que tot procés purament estocàstic (és a dir, sense components deterministes) es pot representar en forma de mitjana mòbil infinita:

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{3.49}$$

on el procés ε_t és soroll blanc i els coeficients ψ_j compleixen que $\sum \psi_j^2 < \infty$ per garantir l'estacionarietat. Noti's que es tracta d'una condició que garanteix que els coeficients siguin prou petits per posar límits a la memòria del procés.

Però fins i tot pel cas no estacionari creiem que és interessant interpretar X_t a partir del model (4.49) sempre tenint en compte que els coeficients ψ_j no tenen perquè complir cap tipus de restricció. Per exemple, si X_t és un camí aleatori, tots els coeficients ψ_j són iguals a la unitat, fet que mostra com els shocks que afecten al valor de la variable són permanents i per tant, té "memòria infinita".

Per tant, sembla clar que una bona manera de caracteritzar la memòria d'un procés estocàstic és a partir d'aquest coeficients que de fet, són anomenats sovint com funció de memòria, ja que no cal fer cap tipus de supòsit inicial sobre l'estacionarietat o ergodicitat de la sèrie.

[18] El supòsit de linealitat està motivat per l'interès d'obtenir models que siguin tractables amb les eines habituals de l'Estadística. La disponibilitat d'eines més sofisticades han permès la introducció durant els darrers anys de gran nombre de models no lineals. De totes maneres, la tesi es centra en els models lineals amb la intenció de no introduïr complicacions formals addicionals i entenent que aquest supòsit no afecta a l'objectiu principal.

Doncs bé, és interessant calcular com són aquests coeficients per cada un dels quatre models plantejats. Els dos primers, corresponen òbviament a models MA finits mentre que el decreixement exponencial correspon a un model AR d'ordre 1 amb coeficient $\phi=\gamma$. Per tant, combinant models ARMA(p,q) es poden aconseguir correlogrames més complexes, per exemple, incloent efectes estacionals però asimptòticament es manté un decreixement exponencial.

Al capítol següent s'introdueix un model senzill que genera un decreixement hiperbòlic però en tot cas, aquests correlogrames no són més que quatre exemples relativament senzills però que permeten il·lustrar com la utilització de models ARMA o, fins i tot, models com els presentats al capítol següent estan subjectes a la crítica que són simplificacions extremes. L'avantatge d'usar aquests models es troba en que permeten una representació senzilla en el domini temporal mentre que altres correlogrames no gaudeixen d'aquesta possibilitat.

Però cal destacar que la única justificació per usar models ARMA o ARFIMA es troba en la constatació empírica que els correlogrames de les dades econòmiques s'hi assemblen prou. Però de fet, amb això no n'hi ha prou perquè hi ha característiques de les dades temporals que no poden observar-se en el domini temporal sinó únicament en el domini de les freqüències. I és en aquest domini on acostumen a observar-se més diferències entre els models ARMA teòrics usats i els espectres empírics o periodogrames.

En aquest sentit, el problema de la memòria llarga analitzat en aquesta tesi no és més que una forma concreta de mostrar les limitacions dels models ARMA. De fet, una de les principals conclusions d'aquesta tesi ha de ser la crítica a l'ús abusiu i sistemàtic dels models ARMA i ARIMA en condicions en que potser aquests no reflecteixen prou bé les característiques de les dades econòmiques i poden portar a interpretacions errònies sobre dites variables.

Noti's que, en tot cas, moltes de les conclusions esmentades al llarg de la tesi són aplicables a altres tipus de dependències estadístiques diferents de l'associada als models ARFIMA.

En primer lloc, i com a etapa prèvia a la definició de la memòria llarga, i donat que els models que contenen arrels unitàries tenen una persistència que no desapareix amb el pas del temps, és a dir, no presenten reversió a la mitjana, sembla raonable distingir aquesta situació que anomenem com de memòria permanent dels casos en que la memòria és transitòria.

Definició 3.1 Es diu que un procés estocàstic X_t presenta memòria transitòria si la correlació entre observacions cada cop més separades en el temps tendeix a zero, és a dir, si:

$$corr(X_t, X_{t+k}) \underset{k\to\infty}{\longrightarrow} 0 \quad \forall t \tag{3.50}$$

Altrament, es diu que el procés presenta memòria permanent.

Aquesta definició és molt general i és més pràctic suposar que els valors observats es poden interpretar en forma de mitjana mòbil infinita, com a l'expressió (3.49). En aquest cas, com:

$$X_{t+k} = X_t + \sum_{i=0}^{k-1} \psi_i \varepsilon_{t+k-i} \tag{3.51}$$

es compleix la següent equivalència:

$$corr(X_t, X_{t+k}) \underset{k\to\infty}{\longrightarrow} 0 \quad \forall t \qquad \Leftrightarrow \qquad \sum_{j=0}^{\infty} \psi_j^2 < \infty \tag{3.52}$$

Per tant, sota la hipòtesi de Yule de representació mitjana mòbil infinita, l'existència de memòria permanent implica la no estacionarietat de les dades i per tant, cal aplicar alguna transformació. Aquesta situació recorda molt l'exemple típic del camí aleatori en que els coeficients mitjana mòbil són constants, $\psi_j=1$. Però aquest no és l'únic model que presenta memòria permanent sinó que un model ARIMA(p,1,q) tampoc compleix la condició (3.52).

De totes maneres, no cal que la representació lineal del procés contingui una arrel unitària per a que no es compleixi la condició (3.52) i per tant tingui memòria permanent[19], però sempre es tractarà de sèries no estacionàries.

[19] Al capítol 4 de la tesi es presenten altres models amb memòria permanent que no contenen arrels unitàries.

També cal esmentar, tot i que al llarg de la tesi es suposa la no presència d'elements deterministes, que la presència d'un tendència determinista també dóna lloc a processos amb memòria permanent.

En particular, els models ARMA tenen tots memòria finita o transitòria però tal com ja s'ha comentat la condició d'estacionarietat o la condició (3.52) no impliquen forçosament que els coeficients d'autocorrelació decreixin de manera exponencial. Noti's que fins aquí s'han definit els diferents conceptes en el domini temporal, però per anar més enllà cal usar el domini de les freqüències. Això és degut a que el domini temporal no és adequat per diferenciar el tipus de memòria quan aquesta és transitòria.

Així, i a partir del que s'ha analitzat al llarg del capítol, sembla raonable proposar la següent definició de memòria llarga[20]:

Definició 3.2. Es diu que un procés estocàstic estacionari de segon ordre amb densitat espectral, presenta memòria llarga si per alguna freqüència $\omega \in [0, \pi]$ l'espectre $f(\omega)$ no està acotat.[21]

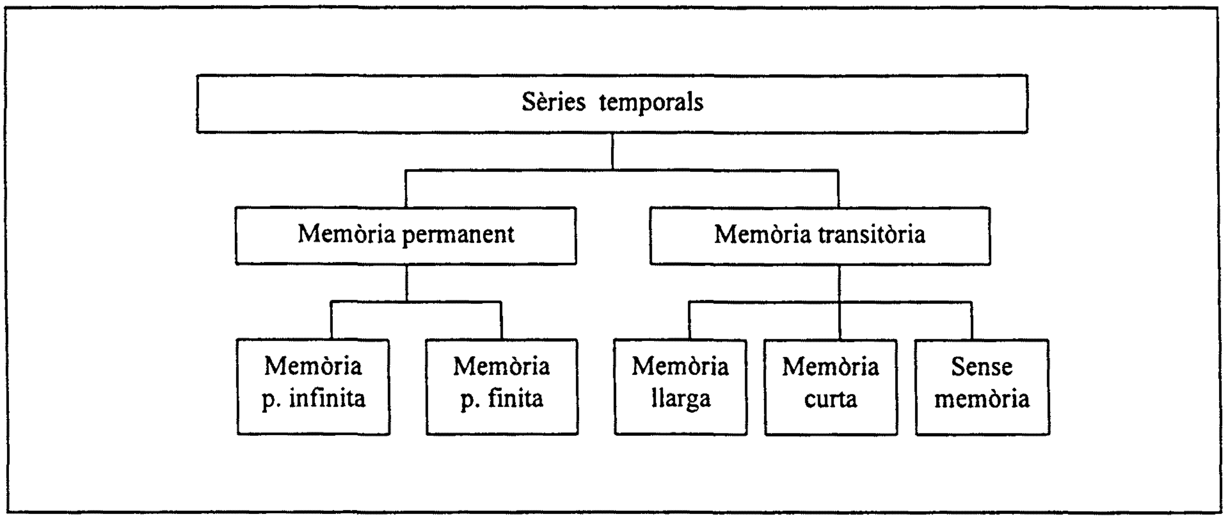
Per exclusió, els processos amb memòria curta són aquells que tenen una densitat espectral continua i acotada:

Definició 3.3. Es diu que un procés estocàstic estacionari de segon ordre amb densitat espectral, presenta memòria curta si l'espectre $f(\omega)$ està acotat per qualsevol freqüència $\omega \in [0, \pi]$.

[20] Tot i que entre els processos amb memòria transitòria, pot haver-n'hi que presentin no estacionarietat diferents de la no estacionarietat en variància, donada la dificultat per tractar amb aquests casos es limita l'anàlisi a l'estacionarietat de segon ordre.

[21] Noti's també com s'exclouen de la definició els processos sense densitat espectral, que considerem com casos clarament patològics i de poc interès.

Quadre 3.1. Classificació de les sèries temporals segons el tipus de *memòria*.



De fet, és aquesta situació la que justifica, formalment, la utilització de models ARMA per aproximar aquests processos. Entre la gran varietat de resultats sobre aquesta possibilitat d'aproximació, es poden citar els següents corol·laris[22]:

Corol·lari 3.1. Si $f(\omega)$ és una densitat espectral continua i $\varepsilon>0$, llavors existeix un model MA(q), $X_t=\theta(L)u_t$ tal que:

$$|f_X(\omega)-f(\omega)|<\varepsilon$$

per qualsevol freqüència $\omega\in[-\pi,\pi]$.

Corol·lari 3.2. Si $f(\omega)$ és una densitat espectral continua i $\varepsilon>0$, llavors existeix un model AR(p), $\phi(L)X_t=u_t$ tal que:

$$|f_X(\omega)-f(\omega)|<\varepsilon$$

per qualsevol freqüència $\omega\in[-\pi,\pi]$.

[22] Recollits a Brockwell i Davis (1992) com Corol·laris 4.4.1 i 4.4.2.

Òbviament, aquests són dos corol·laris de mínims, de manera que en molts casos concrets, serà millor usar un model ARMA que models AR o MA per reduir el nombre de paràmetres seguint el principi de parsimònia.

De fet, es pot demostrar de manera rigorosa que els problemes dels models ARMA(p,q) per reproduir certs comportaments, tot i ser estacionaris, estan directament relacionats amb la presència de singularitats en la densitat espectral. Aquest és el motiu que justifica l'elecció presa en aquesta tesi sobre la definició formal del concepte de memòria llarga. En aquest sentit és destacable que es tracta d'una definició molt enfocada a l'aplicació pràctica de tots aquests conceptes, motivada tal com ja s'ha destacat a la Introducció de la tesi per la preocupació sobre la utilització de models ARMA (o en tot cas, ARIMA, en el cas no estacionari) en casos en que pugui no ser massa correcte.

També és útil per la utilització pràctica d'aquests conceptes, distingir el tipus de memòria llarga de les dades segons la freqüència a la que està associada aquesta forta dependència. Així, es poden distingir alguns casos especialment interessants i habituals en dades econòmiques:

Definició 3.4. Es diu que un procés estocàstic estacionari de segon ordre amb densitat espectral, presenta memòria llarga regular si aquesta funció $f(\omega)$ no està acotada a la freqüència zero.

Definició 3.5. Es diu que un procés estocàstic estacionari de segon ordre amb densitat espectral, presenta memòria llarga regular si aquesta funció $f(\omega)$ no està acotada a alguna de les freqüències estacionals ω_0, és a dir, $\omega_0=2\pi j/s$ per algun $j=1,2...[s/2]$, on s és el nombre d'observacions de la sèrie en un període d'un any.

Naturalment, a nivell teòric es poden presentar altres tipus de memòria llarga, però en tot cas, aquests seran molt poc habituals en dades econòmiques perquè significaria la presència d'un comportament cíclic d'una gran recurrència.

Recordi's que s'ha presentat la noció de memòria llarga a partir de les característiques del correlograma però finalment s'ha optat per donar-ne una definició a partir de la densitat espectral. Per tant, cal preguntar-se quina relació hi ha entre la Definició 3.2 i el comportament de les autocorrelacions. Òbviament, hi ha relació, ja que en una sèrie amb memòria llarga la densitat espectral està relacionada amb les autocovariàncies a través de la següent igualtat:

$$f(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k e^{-ik\omega}=\frac{1}{2\pi}\gamma_o+\frac{1}{\pi}\sum_{k=1}^{\infty}\gamma_k\cos(k\omega) \qquad (3.52)$$

Suposi's que X_t es una sèrie de memòria llarga de manera que la densitat espectral presenta una discontinuïtat en una certa freqüència que ω_0 (que, habitualment serà la freqüència zero). De (3.52) es dedueix que la densitat espectral, $f(\omega)$, és una sèrie de Fourier i per tant, tot i la discontinuïtat compleix bones propietats de regularitat, entre altres que és analítica. De fet, a partir de la densitat espectral és possible obtenir les autocovariàncies com:

$$\gamma_k = \int_{(-\pi,\pi]} f(\omega)e^{ik\omega}d\omega \qquad (3.53)$$

però amb les hipòtesis formulades sobre la densitat no n'hi ha prou per conèixer les propietats de γ_k. Per poder dir-ne alguna cosa cal afegir els següents supòsits tècnics addicionals:

1) Que la densitat espectral $f(\omega)$ és una funció de variació regular[23] l'entorn d'ω_0 i que es pot representar com $f(\omega)=|\omega-\omega_0|^{-2\delta}L(\omega)$ amb $0<\delta<1/4$.[24]
2) Que $L(\omega)\geq 0$ i és de variació acotada als intervals $(0,\omega_0-\varepsilon)\cup(\omega_0+\varepsilon,\pi)$ per qualsevol $\varepsilon>0$.

[23] Per una definició del concepte de funció de variació regular vegi's la nota a peu de pàgina número 29.

[24] En el treball original, s'usa com notació $f(\omega)=|\omega-\omega_0|^{-\beta}L(\omega)$ amb $0<\beta<1/2$ però per coherència amb la notació de l'apartat 3.4 de la tesi, s'usa una notació diferent. Vegi's al respecte també la nota a peu de pàgina nº 8 d'aquest capítol.

Sota aquestes hipòtesis, el Lema de Gray *et al.* (1989) permet demostrar que[25]:

$$\gamma_k \underset{k\to\infty}{\sim} k^{2\delta-1} g(k) \tag{3.54}$$

on[26]:

$$\begin{aligned} g(k) = \cos(k\omega_0)\sin(\pi\delta)\left[L\left(\frac{1}{k}+\omega_0\right)+L\left(\frac{1}{k}-\omega_0\right)\right]- \\ -\sin(k\omega_0)\cos(\pi\delta)\left[L\left(\frac{1}{k}+\omega_0\right)-L\left(\frac{1}{k}-\omega_0\right)\right] \end{aligned} \tag{3.55}$$

Per contra, si l'espectre està acotat en totes les freqüències, es diu que el procés té memòria curta. Noti's que pel fet de definir la memòria llarga sols per processos estacionaris, es compleix que $\sum_{j=0}^{\infty}\gamma_j^2<\infty$ de manera que la correlació entre dades que estan cada cop més distants en el temps tendeix a zero. De manera que en processos amb memòria llarga aquesta memòria acaba per desaparèixer.

Noti's també que el Lema anterior permet relacionar la definició de memòria llarga proposada a la tesi amb les definicions presentades als apartats 3.2 i 3.4

És important, per evitar confusions amb altres treballs en que es manegen conceptes com persistència o memòria, deixar ben clar que s'ha proposat una definició de la memòria com a característica no paramètrica de les sèries temporals (ja siguin o no estacionàries). És important destacar aquest aspecte, perquè alguns dels models analitzats als capítols següents per ajudar en l'anàlisi de les dades temporals com complement dels models ARMA, porten associats certs paràmetres relacionats amb el grau de memòria que sovint es confonen amb el propi concepte de memòria.

---

[25] Vegi's també Gray *et al.* (1994) per la correcció d'un error en l'article inicial.

[26] És a dir, que $\lim_{k\to\infty}\frac{\gamma_k}{k^{2\delta-1}g(k)}=C$ on C és una constant no nul·la.

En aquesta tesi es proposa, en tot cas, distingir clarament el concepte com característica no paramètrica segons les definicions donades de possibles mesures que es puguin proposar, com de fet es proposen posteriorment, del grau de memòria. En tot cas, és important tenir clar que aquestes mesures poden no ser adequades per tots els tipus de memòria, sinó que potser cal usar mesures diferents en cada una de les situacions perfilades.

Per exemple, la via natural de mesurar aquesta memòria en el cas estacionari sembla que ha de ser la representació de Wold, anomenada sovint com funció de memòria. Però aquesta via de mesura sols serveix per la categoria de memòria transitòria estacionària.

3.8. Consideracions finals

L'objectiu d'aquest capítol era el de proposar una classificació de les sèries temporals segons el grau de memòria. A través d'una breu revisió d'algunes de les definicions que es poden trobar a la literatura relacionada amb les sèries temporals, es pot comprovar que tot i que a grans trets, totes aquestes definicions es basen en els mateixos principis abstractes, un cop es presenten definicions formals aquestes no són coincidents.

Les definicions proposades a l'apartat 3.7 són perfectament discutibles però creiem que són les més correctes per a l'anàlisi economètrica. Aquest és un aspecte a destacar ja que és molt probable que en altres àmbits de l'anàlisi estadística pugui ser interessant usar altres classificacions però creiem que la proposada és adequada per abordar a la tercera part de la tesi l'anàlisi de les conseqüències que té el tipus de memòria a l'anàlisi economètrica.

Pel que fa a la classificació proposada insistir en les que, al nostre entendre, són les novetats principals d'aquesta classificació:

- D'una banda, es tracta d'una classificació que combina unes definicions basades en els domini temporal amb altres basades en el domini de les freqüències.

- Pel que fa a la distinció entre memòria curta i memòria llarga es pot trobar la mateixa definició en altres treballs tot i que no n'és la més popular a la literatura economètrica. S'ha optat per aquesta definició per ser la més general.

- La distinció entre memòria permanent i memòria transitòria tot i no ser equivalent es pot interpretar, de manera simplificada com la distinció habitual entre estacionarietat i no estacionarietat.

Atès que els models ARMA són adequats per a la representació de les sèries amb memòria curta, aquest tipus de memòria no presenta cap problemàtica especial. En canvi, sí que mereix especial atenció les conseqüències de la presència de memòria llarga. Com a eina prèvia i necessària per aquesta anàlisi es dedica el capítol següent a la presentació d'alguns models teòrics que poden usar-se per representar sèries amb memòria llarga.

En tot cas, i pel que fa als diferents tipus de sèries amb memòria permanent, com són no estacionàries cal deixar ben clar que per a la seva representació caldrà usar algun tipus de transformació prèvia per convertir les dades en estacionàries.

SEGONA PART

MODELITZACIÓ I INFERÈNCIA

Capítol 4

MODELITZACIÓ DE SÈRIES ECONÒMIQUES AMB MEMÒRIA LLARGA

4.1. Introducció

A continuació, un cop s'han definit de manera formal els diferent tipus de memòria, cal disposar de models adequats que permetin aproximar el comportament de les variables econòmiques segons el tipus de memòria que presenten.

Així, suposi's que es disposa de T observacions d'una certa variable econòmica, $X_1,..,X_T$. Si és estacionària, un cop eliminats els possibles components deterministes, pel Teorema de Wold es pot representar en forma de mitjana mòbil infinita $X_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k}$ on ε_t és soroll blanc.

Si la densitat espectral de la sèrie està acotada[1] i per tant sols conté memòria curta, té sentit aproximar el seu comportament identificant i estimant un model ARMA, però cal disposar d'algun instrument per modelitzar aquesta sèrie quan això no sigui així. Aquest capítol es centra en l'anàlisi de diferents alternatives per la modelització de sèries amb memòria llarga.

A partir del moment en que s'obté evidència empírica de dades econòmiques que contenen una persistència o memòria important, neix l'interès pel desenvolupament de models capaços de reproduir aquest tipus de comportament.

Així, es genera en alguns autors la concepció[2], defensada a la tesi, que per modelitzar unes certes dades temporals cal usar un model amb unes característiques similars pel que fa a memòria a les que presenten les dades que són objecte d'estudi.

De totes maneres, quantes més característiques concretes s'exigeixin als models usats, més complexes han de ser aquests models. De manera que sempre cal valorar els beneficis que pot

[1] Noti's que resta pendent de capítols posteriors la recerca d'algún mètode que permeti identificar el tipus de memòria que contenen les dades.

[2] Defensada per exemple a Parzen (1983).

reportar l'extensió del tipus de models usats. Aquest és un aspecte molt important pel que fa a la modelització de la memòria llarga.

Els primers models que es desenvolupen usen la notació introduïda per Hurst (1951) i estan, d'alguna o altra manera, molt associats a l'estadístic R/S presentat al capítol anterior. D'altra banda, els primers models suggerits són a temps continu tot i que posteriorment es desenvolupen més models a temps discret.

Val la pena assenyalar també que els primers intents de reproduir aquest tipus de comportament en les dades econòmiques estan molt lligats a Mandelbrot que com s'ha comentat més amunt era molt crític amb la utilització de les correlacions i el periodograma per analitzar dades temporals econòmiques.

Posteriorment esdevenen populars a la metodologia economètrica els models ARMA, de manera que no és estrany que es proposin models de memòria llarga que són una extensió d'aquest tipus de modelització.

Així, la segona onada de models, ja a la dècada dels anys vuitanta està molt lligada als treballs de Robinson (1978) i Granger (1980) en que es mostra com per efecte de l'agregació sorgeix de manera natural el fenomen de la memòria llarga.

Aquesta segona època està molt associada als models ARFIMA a partir dels treball de Granger i Joyeaux (1980) i Hosking (1981). Cal també destacar els intents recents d'establir una teoria més general, basada en nocions no paramètriques o semiparamètriques de la densitat espectral.

A continuació, es presenten algunes d'aquestes propostes per modelitzar de manera explícita la presència de memòria llarga. Cal destacar com a última d'aquestes modelitzacions, una possibilitat per la que, sota el nostre punt de vista, s'ha mostrat poc interès fins el moment. Es tracta d'una extensió del model exponencial proposat per Bloomfield (1973) al que es dedica l'últim apartat d'aquest capítol i que s'utilitza després al llarg de la tesi.

També cal destacar que l'objecte d'aquest capítol és exclusivament l'anàlisi de les possibles aproximacions paramètriques a la modelització de la memòria llarga deixant per a capítols posteriors el tractament no paramètric o semiparamètric de la memòria ja que aquest tractament no permet obtenir models explícits tot i que poden significar vies de coneixement de les característiques de les dades molt interessants.

A més, es posa especial interès a presentar breument les característiques principals de cada un d'aquests models per tal d'identificar en quines situacions pot ser interessant la seva utilització.

4.2. El moviment brownià fraccional

La primera proposta de modelització explícita-de memòria llarga es troba a Hurst (1951) que associa aquest comportament amb la presència d'unes autocorrelacions del tipus següent:

$$\gamma_k \sim Ak^{2(H-1)} \tag{4.1}$$

Segons aquest autor, el coeficient H mesura el grau de memòria i defineix el concepte de memòria llarga com aquella situació en que $H>1/2$. Aquest coeficient s'ha popularitzat amb el nom de *coeficient de Hurst* i a la presència de valors d'H superiors a $1/2$ se la coneix amb el nom d'efecte Hurst.

Aquesta definició, en tot cas, està inclosa en la definició de memòria llarga adoptada en aquesta tesi ja que, si un procés estacionari compleix la condició (4.1), té una densitat espectral no acotada i, per tant, segons la definició 3.2 és de memòria llarga. En efecte, la condició (4.1) significa que existeix una constant C de manera que $\lim_{k\to\infty} \frac{\gamma_k}{k^{2(H-1)}} = C$. Per tant, si $H>1/2$ es compleix que $2(H-1)>1$ i la sèrie $\sum_{k=1}^{\infty} \gamma_k$ és divergent. Per tant, la densitat espectral avaluada a la freqüència zero no està acotada:

$$f(0)=\frac{1}{2\pi}\gamma_o+\frac{1}{\pi}\sum_{k=1}^{\infty}\gamma_k=\infty$$

De fet, la definició de Hurst (1951) és un cas particular de la definició *F2* de l'apartat 3.7 on l'Índex del grau de memòria és $\delta=H+1/2$.

De totes maneres, Hurst (1951) no proporciona un model explícit que presenti unes autocorrelacions que compleixin la condició (4.1). De totes maneres, alguns anys després, a Mandelbrot (1965) i Mandelbrot i Ness (1968) es presenta el primer model explícit que té unes autocorrelacions d'aquest tipus. Seguint aquests treballs, es defineix el moviment brownià fraccional com una generalització del moviment brownià estàndard[3]:

Definició 4.1. Es diu que un procés estocàstic a temps continu $W_H(t)$ és un moviment brownià fraccional amb $H\in(0,1)$ si:

$$W_H(r)=\frac{1}{\Gamma(H+1/2)}\int_0^r(r-s)^{H-1/2}dW(s) \tag{4.2}$$

on $\Gamma(\cdot)$ és la funció gamma i $W(s)$ és un moviment brownià estàndard.

Entre les propietats més destacables d'aquest procés estocàstic cal esmentar les següents:

1) Existeix per 0<H<1 i és quasi segurament continu
2) És la derivada fraccional *(1/2-H)*-èssima d'un procés brownià[4].
3) Té densitat espectral proporcional a ω^{-2H-1}.
4) Quan $H=1/2$ s'obté un procés brownià estàndard.
5) Per un valor r fixat s'obté:

$$W_H(r)\sim N\left(0,\frac{r^{2H}}{\Gamma(H+1/2)^2\Gamma(2H)}\right) \tag{4.3}$$

[3] Vegi's Hamilton (1994) per una introducció senzilla al concepte de procés brownià.

[4] Derivada fraccional en el sentit de Weyl o Riemann-Liouville.

6) Pel que fa a la funció d'autocovariàncies és proporcional a $|k|^{2H-2}$

7) Així com el soroll blanc a temps continu és la derivada d'un procés brownià, es pot definir també el concepte de procés soroll blanc fraccional a temps continu com la derivada d'un moviment brownià fraccional, és a dir, $W'_H(t)$, o bé com la $(1/2-H)$-èssima derivada fraccional d'un soroll blanc a temps continu.

Tot i que el moviment brownià permet modelitzar sèries econòmiques amb memòria llarga, es tracta d'un procés estocàstic a temps continu i per simplicitat, es restringirà la tesi a l'anàlisi de processos estocàstics a temps discret. De totes maneres, el moviment brownià fraccional és important per dos aspectes:

- perquè és històricament el primer procés estocàstic amb memòria llarga, i
- perquè és l'antecedent directe d'altres processos estocàstics a temps discret que es presenten a continuació i que també presenten memòria llarga.

4.3. El soroll gaussià fraccional

El primer model a temps discret que pretén recollir la característica que hem definit com memòria llarga fou introduït també per Mandelbrot i Wallis (1969) com una versió discreta del moviment brownià fraccional que consisteix a definir un procés a temps discret amb covariàncies iguals a les d'una diferència de moviments brownians fraccionals, és a dir, $W_H(t)-W_H(t-1)$.

Aquest concepte està molt associat al utilitzat prèviament per Kolmogorov (1940) sota el nom de procés estocàstic *self-similar*. Amb aquest nom nota que en un procés estocàstic a temps continu es compleix:

$$F(X(t_1), \ldots, X(t_n)) = a^{-H} F(X(at_1), \ldots, X(at_n)) \tag{4.4}$$

De fet, la propietat de *self-similar* està associada amb el fet que la forma de la densitat espectral no depengui de l'interval amb que es mesuren les dades. A alguns estudis, la persistència es considera relacionada amb l'interval amb que s'observen les dades. En la terminologia de Mandelbrot (1972, 1975) un *fractal* es defineix com:

$$|\Delta X_t| = (\Delta t/c)^H \tag{4.5}$$

on c és una constant. En aquesta terminologia, un fractal és *self-similar* si té el mateix H per a qualsevol elecció de l'interval d'observació.

Definició 4.2. Sigui Y_t un procés a temps discret definit per:

$$Y_t = \int_{-\infty}^{t} (t-s)^{H-1/2}\, dW(s) \tag{4.6}$$

on $W(s)$ és un procés Brownià estàndard i $H \in (0,1)$, llavors es defineix $X_t = Y_t - Y_{t-1}$ com un soroll gaussià fraccional.

Jonas (1981) demostra que la densitat espectral d'un soroll gaussià fraccional és[5]:

$$f(\omega;H) = \frac{\sigma^2}{(2\pi)^{-2H-2}} \Gamma(2H+1)\sin(\pi H) 4\sin^2(\omega/2) \sum_{n=-\infty}^{\infty} |n + \omega/2\pi|^{-2H-1} \tag{4.7}$$

És fàcil comprovar com al voltant de la freqüència zero es compleix:

$$\lim_{\omega \to \infty} \frac{f(\omega,H)}{\omega^{2H-1}} = \frac{2\sigma^2}{\pi} \Gamma(2H+1)\sin(\pi H) \tag{4.8}$$

de manera que l'espectre és finit a l'origen sols si $H \leq 1/2$ mentre que si $H > 1/2$ la sèrie presenta memòria llarga.

La funció d'autocorrelació és:

$$\rho_k = \frac{1}{2}\left\{ |k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H} \right\} \tag{4.9}$$

[5] On s'usa el conveni que la densitat espectral és, per definició $f(\omega) = \sum \gamma_k \exp(-ik\omega)$.

de manera que per retards grans:

$$\rho_k \sim H(2H-1)k^{2H-2} \tag{4.10}$$

Aquest model admet una generalització immediata a partir d'admetre a més una certa estructura de dependència amb memòria curta.

Definició 4.3. Un soroll gaussià fraccional general és un procés estocàstic amb densitat espectral $f(\omega)=f_U(\omega)f(\omega;H)$ on $f_U(\omega)$ és una funció continua, acotada i no negativa en l'interval $[-\pi,\pi]$.

En particular, es pot tractar de la densitat espectral d'un procés ARMA.

A més, a Mandelbrot i Taqqu (1979) es demostra que el soroll gaussià fraccional és l'únic procés gaussià tal que $S_{mT}=\sum_{t=1}^{mT}X_t$ té la mateixa distribució que $m^H S_T$ $\forall m,T\geq 1$.

De totes maneres, es tracta d'un procés estocàstic difícil de tractar formalment, sobretot si se'l compara amb el soroll blanc fraccionalment integrat i els models ARFIMA que es presenten a continuació. A més, tal com es comenta al següent apartat, hi ha una forta relació entre el soroll gaussià fraccional i el soroll blanc fraccionalment integrat.

4.4. El soroll blanc fraccionalment integrat (SBFI)

Mentre que el cas discret anàleg al moviment brownià és el camí aleatori, la versió discreta del moviment brownià fraccional és el soroll blanc fraccionalment integrat en que es substitueix l'operador diferenciació del camí aleatori per l'operador diferència fraccional $(1-L)^d$ on d no té perquè ser un nombre natural. Aquest operador es defineix, quan $d>-1$ a través de la seva expansió binomial:

$$(1-L)^d=\sum_{k=0}^{\infty}\binom{d}{k}(-L)^k=1-dL-\frac{1}{2}d(1-d)L^2-\frac{1}{6}d(1-d)(2-d)L^3\ldots \tag{4.11}$$

Definició 4.4. Un procés X_t sense components deterministes es diu que és un *SBFI(d)* o soroll blanc fraccionalment integrat amb $d \in (-0.5, 0.5)$ si X_t és una solució estacionària de la següent equació estocàstica en diferències:

$$(1-L)^d X_t = \varepsilon_t \tag{4.12}$$

on ε_t és soroll blanc.

A Brockwell i Davis (1987) es demostra que per $d \in (-0.5, 0.5)$ existeix una única solució estacionària d'aquesta equació.

Tot i que el procés definit té esperança zero, es pot generalitzar amb la inclusió d'elements deterministes i definir X_t a partir de $(1-L)^d (X_t - \mu) = \varepsilon_t$ de manera que $E(X_t)=\mu$ o bé a partir de $(1-L)^d (X_t - \mu - \beta t) = \varepsilon_t$ de manera que $E(X_t)=\mu+\beta t$. A partir d'aquest moment, i sense pèrdua de generalitat es suposa esperança zero. En cas contrari n'hi ha prou amb centrar les dades.

A Hosking (1981) es demostra que per $d>-1$ es tracta d'un procés invertible de manera que si X_t és SBFI(d) es pot representar en forma autoregressiva:

$$X_t = \varepsilon_t + \sum_{k=1}^{\infty} \xi_k X_{t-k} \tag{4.13}$$

A partir de l'expansió binomial (4.11), s'obté que aquests coeficients autoregressius són iguals a:

$$\xi_k = \frac{-d(1-d)....(k-1-d)}{k!} = \prod_{j=0}^{k} \frac{j-1-d}{d} = \frac{(k-d-1)!}{k!(-d-1)!} \tag{4.14}$$

Usant la funció gamma és possible obtenir formulacions alternatives d'aquests coeficients i que a efectes pràctics són força útils. Recordi's que la funció gamma, $\Gamma(x)$ es defineix, per $x>0$ com:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \tag{4.15}$$

Quan x és un valor no positiu és habitual usar una extensió d'aquesta funció de manera que $\Gamma(x)=x^{-1}\Gamma(x+1)$ quan $x<0$ i $\Gamma(0)=\infty$.

Una de les propietats bàsiques de la funció gamma és:

$$\Gamma(k-d)=(k-d-1)(k-d-2)..(2-d)(1-d)(-d)\Gamma(-d) \tag{4.16}$$

i a partir d'aquesta propietat, es pot simplificar l'expressió (4.14) de manera que els coeficients ξ_k es poden calcular com:

$$\xi_k=\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} \tag{4.17}$$

Una altra manera d'expressar els coeficients ξ_k és a través de la funció hipergeomètrica $F(a,b;c;x)$ de manera que:

$$(1-L)^d=\sum_{k=0}^{\infty}\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)}L^k=F(-d,1,1;L) \tag{4.18}$$

Recordi's que la funció hipergeomètrica $F(a,b;c;x)$ es defineix com[6]:

$$\begin{aligned}F(a,b;c;x)&=1+\frac{a\times b}{c\times 1}x+\frac{a\times(a+1)\times b\times(b+1)}{c\times(c+1)\times 1\times 2}x^2+\ldots=\\&=\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{j=1}^{\infty}\frac{\Gamma(a+i)\Gamma(b+j)}{\Gamma(c+j)\Gamma(j+1)}x^j\end{aligned} \tag{4.19}$$

De totes maneres, la majoria de vegades que hom necessita calcular els coeficients ξ_k és molt més senzill fer-ho a través d'un mètode recursiu. Concretament, de l'expressió (4.14) es dedueix que es compleix la següent condició:

---

[6] A efectes pràctics, la funció hipergeomètrica es pot calcular o bé a través d'algún algorisme per aproximar la funció gamma o bé de manera exacta a partir de la següent expressió:

$$F(a,b;c;x)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\left[I(a,b;c;x)+(1-x)^{-a}I\left(a,c-b;c;\frac{x}{x-1}\right)\right]$$

on:

$$I(a,b;c;x)=\frac{(1/2)^b}{b}+\int_0^{1/2}t^{b-1}\left[(1-t)^{c-b-1}(1-tx)^{-a}-1\right]dt$$

$$\xi_k = \xi_{k-1} \frac{k-1-d}{k} \tag{4.20}$$

Com a més es compleix:

$$\xi_0 = 1 \tag{4.21}$$

aquestes dos propietats permeten calcular els coeficients ξ_k de manera molt simple. A més, a partir de la propietat (4.20) es pot comprovar que asimptòticament aquests coeficients compleixen la següent aproximació:

$$\xi_k \underset{k\to\infty}{\sim} \frac{k^{-d-1}}{(-d-1)!} \tag{4.22}$$

Per tant, tot i que els coeficients ξ_k tendeixen cap a zero a mesura que k tendeix cap a infinit, no existeix un retard a partir del qual tots els coeficients autoregressius siguin zero. De fet, de (4.22) es dedueix que aquests coeficients tendeixen molt lentament cap a zero.

D'aquestes observacions es dedueix que es poden interpretar els processos SBFI(d) com casos particulars de processos AR(∞) amb coeficients que compleixen la condició (4.20). El problema que cal plantejar, arribats a aquest punt és la possible existència de bones aproximacions AR finites que permetin simplificar el tractament d'aquest tipus de comportament.

Doncs bé, tot i que la successió de coeficients autoregressius del model SBFI(d) tendeixin cap a zero, el tret fonamental d'aquest model és precisament l'existència d'aquests coeficients autoregressius diferents de zero per retards grans, independentment de la magnitud d'aquests coeficients. Aquest és justament l'origen de que hi hagi una gran persistència en aquestes dades.

Aquesta reflexió permet intuir els greus problemes que han de patir moltes de les tècniques estadístiques més habituals a l'hora de detectar aquesta successió de coeficients ja que valors autoregressius estimats propers a zero porten a interpretar en termes d'inferència que els valors poblacionals són nuls. En tot cas, aquests problemes en la detecció de la memòria llarga són objecte d'ampli estudi posteriorment.

Un cop analitzada la representació autoregressiva dels models SBFI(d), també es pot demostrar l'existència d'una representació en forma de mitjana mòbil infinita. De fet, si X_t és SBFI(d) amb $d<1$, com és un procés invertible, es pot representar en forma de mitjana mòbil infinita:

$$X_t = \sum_{k=1}^{\infty} \psi_k \varepsilon_{t-k} \tag{4.23}$$

Per trobar el valor dels coeficients ψ_k n'hi ha prou amb partir de la base de que X_t compleix l'equació $(1-L)^d X_t=\varepsilon_t$. Per tant, si s'inverteix aquesta expressió s'obté que $X_t=(1-L)^{-d}\varepsilon_t$, de manera que, usant (4.14) amb $-d$ enlloc de d, s'obté que els coeficients ψ_k es poden calcular com:

$$\psi_k = \frac{d(1+d)\ldots(k-1+d)}{k!} = \prod_{j=0}^{k} \frac{j-1+d}{j} = \frac{(k+d-1)!}{k!(d-1)!} \tag{4.24}$$

Novament, es pot usar la funció gamma per donar una expressió més simple d'aquests coeficients. Concretament, com la funció gamma compleix:

$$\Gamma(d+k) = \frac{d(d+1)(d+2)..(d+k+1)}{\Gamma(d)} \tag{4.25}$$

s'obté que els coeficients mitjana mòbil que apareixen a (4.24) es poden expressar també com:

$$\psi_k = \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} \tag{4.26}$$

De manera similar al cas autoregressiu, també és possible trobar una expressió recursiva que permeti calcular de manera còmode aquests coeficients. Concretament:

$$\psi_k = \psi_{k-1} \frac{k-1+d}{k} \tag{4.27}$$

Aquesta coeficients també tendeixen a zero a mesura que el retard tendeix a infinit ja que asimptòticament es compleix que:

$$\psi_k \underset{k\to\infty}{\sim} \frac{k^{d-1}}{(d-1)!} \tag{4.28}$$

Recordi's que el teorema de Wold garanteix aquesta representació per qualsevol procés estacionari sense part determinista. A més, per la definició 3.1, un procés mitjana mòbil infinit conté memòria transitòria o permanent segons el sumatori $\sum_{k=0}^{\infty} \psi_k^2 < \infty$ sigui finit o infinit, respectivament. Així, en el cas d'un soroll fraccionalment integrat, és fàcil comprovar que:

$$\sum_{k=0}^{\infty} \psi_k^2 \sim \sum_{k=0}^{\infty} \frac{k^{2(d-1)}}{\Gamma(d)} < \infty \quad \Leftrightarrow \quad d < \frac{1}{2} \tag{4.29}$$

i per tant, presenta memòria transitòria quan $d<1/2$ i memòria permanent *quan* $d \geq 1/2$.

Donat que s'han definit els conceptes de memòria llarga i memòria curta d'una sèrie temporal a partir de la densitat espectral, té especial interès parar atenció a les característiques de l'espectre d'un model SBFI(d).

Si X_t és SBFI(d) ha de ser solució de l'equació estocàstica $(1-L)^d X_t = \varepsilon_t$. Per tant, si existeix la densitat espectral, la manera més senzilla de calcular aquesta és traduir aquesta equació al domini de les freqüències substituint l'operador L per $exp(-i\omega)$, X_t per $f(\omega)$ i ε_t per $f_\varepsilon(\omega)$. Com ε_t és soroll blanc, la seva densitat espectral és:

$$f_\varepsilon(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \tag{4.30}$$

i s'obté que la densitat espectral d'X_t és:

$$f(\omega) = \frac{\sigma^2}{2\pi}\left[1 - e^{-i\omega}\right]^{-2d} = \frac{\sigma^2}{2\pi}\left[2sin(\omega/2)\right]^{-2d} \tag{4.31}$$

De fet, aquesta expressió sols té sentit com a densitat espectral en sentit estricte quan X_t és estacionari, o sigui per $d<1/2$. Observi's com en aquest cas, al voltant de la freqüència zero es compleix que:

$$f(\omega) \underset{\omega \to 0}{\sim} C\omega^{-2d} \tag{4.32}$$

on C és una constant.

A partir de l'aproximació (4.32) s'observa que quan $0<d<1/2$, la densitat espectral no està acotada ja que tendeix cap a infinit en la freqüència zero. Per tant, atenent a la definició 3.2 un procés SBFI(d) amb $d>0$ presenta memòria llarga. En canvi, quan $d=0$, el procés X_t és senzillament un soroll blanc ja que l'espectre és constat, és a dir, sense memòria.

Per altra banda, quan $d<0$, l'espectre està acotat en totes les freqüències i per tant, d'acord amb la definició que s'ha adoptat en aquesta tesi presenta memòria curta. Observi's també que en aquest cas l'espectre no és constant i per tant el procés no és soroll blanc. En la terminologia de Mandelbrot (1977) aquest procés és *antipersistent* ja que presenta un tipus de persistència negativa.

Per tant, i com a conclusió, els models SBFI(d) són models molt senzills que, tot i la seva simplicitat, permeten la modelització de dades amb memòria llarga. De totes maneres, mentre que la representació de dades econòmiques en el domini temporal a través de l'operador $(1-L)$ té una interpretació natural, la interpretació de la diferenciació fraccional no és fàcil. En tot cas, i a falta de recerca addicional que permeti donar-ne altres interpretacions, s'ha d'interpretar un model SBFI(d) com a un model AR(∞) o MA(∞).

A més de les característiques ja analitzades, és important conèixer també les propietats dels processos SBFI(d) en el domini temporal, és a dir, les autocovariàncies, autocorrelacions i autocorrelacions parcials per comparar-les, per exemple, amb les de models ARMA.

Pel que fa a les autocovariàncies, es demostra a Granger i Joyeaux (1980) i Hosking (1981) que la variància d'un procés X_t és:

$$\gamma_0 = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)} \tag{4.33}$$

A partir de la següent igualtat[7]:

[7] Per una demostració es pot consultar Gradshteyn i Ryzhik (1965).

$$\int_0^{\pi} \cos(hx) sin^{k-1}(x)dx = \frac{\pi \cos(h\pi/2)\Gamma(k+1)2^{1-k}}{k\Gamma((k+h+1)/2)\Gamma((k-h+1))/2)} \tag{4.34}$$

es pot comprovar que per cada retard k les autocovariàncies són iguals a[8]:

$$\begin{aligned}
\gamma_k &= 2\int_0^{\pi} \frac{\sigma^2}{2\pi} \cos(k\omega)\left[sin(\omega/2)\right]^{-2d} d\omega = \\
&= \frac{(-1)^k(-2d)!}{(k-d)!(-k-d)!}\sigma^2 = \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)}\sigma^2 = \\
&= \frac{\sigma^2 sin(\pi d)}{2\pi} \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)}\sigma^2 = (-1)^k \frac{\Gamma(1-2d)}{\Gamma(1-k-d)\Gamma(k+1-d)}\sigma^2
\end{aligned} \tag{4.35}$$

Pel que fa a les autocorrelacions, es pot comprovar com els primers coeficients són:

$$\rho_1 = \frac{d}{1-d} \tag{4.36}$$

i:

$$\rho_2 = \frac{d(1+d)}{(1-d)(2-d)} \tag{4.37}$$

mentre que, en general:

$$\rho_k = \prod_{i=0}^{k} \frac{i-1+d}{i-d} = \frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)} \tag{4.38}$$

Observi's com asimptòticament aquestes correlacions tendeixen a zero ja que:

$$\rho_k \underset{k\to\infty}{\sim} \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1} \tag{4.39}$$

De totes maneres, a efectes pràctics de càlcul és més adequat usar una fórmula recursiva més senzilla:

$$\rho_k = \rho_{k-1} \frac{(d+j-1)}{(j-d)} \tag{4.40}$$

Finalment, aplicant les equacions de Yule-Walker a les autocorrelacions anteriors, s'obtenen les autocorrelacions parcials:

[8] Es presenten algunes expresions equivalents d'aquestes autocovariàncies que seran útils al llarg de la tesi.

$$\phi_{kk} = \frac{d}{k-d} \qquad (4.41)$$

Fins al moment sols s'han considerat els models SBFI(d) amb $d<1/2$ ja que així X_t és estacionari. De totes maneres, és interessant considerar també els casos en que X_t compleix l'equació $(1-L)^d X_t = \varepsilon_t$ amb $d \geq 1/2$.

En aquest cas X_t no és estacionari de manera que no existeix, de manera estricta, la densitat espectral del procés. De totes maneres, es pot estendre la noció de pseudo-espectre dels models integrats a aquest cas[9].

Definició 4.5. El procés X_t centrat es diu que és un SBFI(d) amb $d \geq 0.5$ si existeix un valor d^* natural de manera que $Y_t = (1-L)^{d^*} X_t$ és un procés SBFI(d-d^*) amb $d-d^* \in (-0.5, 0.5)$.

Observi's que en aquest context, un camí aleatori és un cas molt particular d'un model SBFI(d) en que $d=1$. A més, aquesta definició inclou com a cas particular els sorolls blancs integrats d'ordre d en que $d^*=d$.

Els models SBFI(d) són una conseqüència natural de la utilització en el domini temporal de l'operador diferenciació fraccional però aquesta no és la seva única aplicació sinó que permet introduir un conjunt més ampli de models. Més concretament, a partir de les definicions anteriors per diferents valors del paràmetre d, es pot establir una generalització del concepte d'integració. Recordi's que una sèrie temporal es diu que és integrable d'ordre d per valors naturals del paràmetre d si un cop diferenciada la sèrie d cops és estacionària, formalment:

Definició 4.6. Un procés sense termes deterministes X_t és integrable d'ordre d natural, denotat per si $X_t \sim I(d)$ si $U_t = (1-L)^d X_t$ té una densitat espectral continua, acotada i no negativa[10].

[9] Per una definició del concepte de pseudo-espectre vegi's el capítol 11 de Priestley (1981).

[10] Sovint es defineix el concepte de sèrie integrada (amb d natural) imposant que un cop aplicat l'operador diferenciació s'obté un model ARMA. De totes maneres, aquesta pot ser una definició massa restrictiva i sembla

Així, a partir del concepte de procés soroll blanc fraccionalment integrat sembla natural la següent generalització de la definició 4.6:

Definició 4.7. El procés X_t centrat es diu que és integrable d'ordre d on $d>-1/2$ si $U_t = (1-L)^d X_t$ té una densitat espectral continua, acotada i no negativa. En aquest cas, per analogia amb la notació habitual amb ordres d'integració naturals, es diu que X_t és I(d).[11]

Observi's que la definició 4.7 permet una generalització dels processos SBFI(d) ja que U_t no té perquè ser soroll blanc. Això permet que un procés X_t que sigui FI(d) contingui tant components de memòria curta com components de memòria llarga. Així, els models ARFIMA introduïts a l'apartat següent són un cas particular de processos FI(d).

Un altre aspecte a destacar és que aquesta noció d'integració està molt relacionada amb el concepte de soroll gaussià fraccional general de la definició 4.3. De fet, Geweke i Porter-Hudak (1983) en demostren formalment l'equivalència a través del següent teorema:

Teorema 4.1. X_t és un procés integrable d'ordre d si i només si també es pot representar com un soroll gaussià fraccional general amb $H=d-1/2$.

En definitiva, l'operador $(1-L)^d$ amb $d<1/2$ genera unes estructures d'autocorrelacions hiperbòliques, diferents de les clàssiques estructures exponencials i per tant permeten modelitzar la presència de memòria llarga.

natural definir el concepte d'integració imposant que un cop aplicat aquest operador s'obté un procés amb memòria curta.

[11] De vegades és habitual referir-se a processos FI(d) quan es vol destacar que el paràmetre d no té perquè ser un valor natural. En aquests casos és habitual dir que el procés està fraccionalment integrat. A la tesi s'utilitzarà sovint la notació FI(d) per destacar aquest fet.

Però l'operador diferència fraccional té a més una altra utilitat. Suposi's que $(1-L)^d X_t=\varepsilon_t$ amb $d\geq 1/2$. Per la propietat (4.29) conté memòria infinita però a diferència d'un camí aleatori, els shocks no tenen uns efectes permanents. Aquesta és una conclusió molt important cara a la mesura de la persistència que presenten les variables econòmiques.

Com han mostrat Campbell i Mankiw (1987), a partir de la següent representació:

$$(1-L)X_t = A(L)\varepsilon_t \tag{4.42}$$

la persistència dels shocks ve determinada a partir de $A(1)$. Com:

$$A(L) = (1-L)^{1-d} = F(d-1,1;1;L) \tag{4.43}$$

es comprova[12] que $F(d-1,1;1;L)=0$ si $d<1$, i per tant $A(1)=0$. Així, l'efecte d'un shock en t sobre $t+k$ és $1+A_1+A_2+...A_k$. Per a que un procés presenti reversió a la mitjana cal que A(1)=0, és a dir, quan $1/2\leq d<1$, el procés presenta reversió a la mitjana tot i no presentar memòria infinita.

Abans d'estendre el procés SBFI(d) a modelitzacions més complexes a l'apartat següent, cal esmentar dos grans objeccions que es poden fer, al nostre entendre, a la noció d'integració fraccional. D'una banda, l'aspecte ja citat de la falta d'interpretació clara. Però, d'altra banda, al marge de que es tracta d'un tipus de models interessants perquè permeten modelitzar la presència de memòria llarga, fora bo disposar d'alguna justificació addicional per la utilització d'aquest model.

De fet, donat que s'ha definit el concepte de memòria llarga en el domini de les freqüències, sembla raonable valorar la seva modelització també en aquest domini. En aquest sentit, la densitat espectral dels models SBFI(d) no és més que un tipus molt concret entre una gran varietat teòrica de possibilitats. El motiu de que hagi estat objecte de gran interès durant els últims anys, és bàsicament que es pot representar de manera còmode en el domini temporal.

[12] A partir de Gradszteyn i Ryzhnik (1980 pàg. 1039-1042).

A diferència dels models ARMA, que es poden justificar a partir d'un Teorema de representativitat, no es disposa d'algun resultat d'aquest tipus pels models SBFI(d) o pels models ARFIMA analitzats després. En tot cas, es pot destacar que el treball de Granger (1980) que aporta una justificació teòrica a la utilització d'aquests models per a la modelització de dades temporals que provenen de l'agregació de dades individuals.

Concretament, si la variable observada X_t és l'agregació de N variables:

$$X_t = \sum_{j=1}^{N} Y_{jt} \tag{4.44}$$

on cada una d'aquestes variables segueix un procés autoregressiu:

$$Y_{jt} = \rho_j Y_{jt-1} + \varepsilon_{jt} \tag{4.45}$$

llavors, com és ben conegut, la variable X_t segueix, en general, un model ARMA(N,N-1). Granger s'interessa pel comportament de la variable X_t quan el nombre de sèries agregades tendeix cap a infinit. Per fer-ho suposa que els coeficients autoregressius ρ_j estan distribuïts segons una distribució Beta (p,q) a l'interval [0,1]:

$$f(x) = \frac{2}{B(p,q)} x^{2p-1}(1-x^2)^{q-1} \quad x \in [0,1] \tag{4.46}$$

i demostra que en aquest cas, al tendir N cap a infinit, X_t tendeix cap a un model SBFI(1-$q/2$).

4.5. Els models ARFIMA

Una de les limitacions del model SBFI(d) és que, limitant l'anàlisi a $d \geq 0$ que és l'habitual en dades econòmiques, o bé la sèrie presenta memòria llarga (*quan* $d>0$) o bé no presenta memòria (quan $d=0$) eliminant la possibilitat de que contingui sols memòria curta, de manera que fora interessant disposar d'una extensió d'aquests models que permeti incloure situacions intermèdies.

L'extensió més intuïtiva d'aquests models que permet aquesta possibilitat consisteix a utilitzar l'operador diferenciació fraccional per generalitzar no sols el model camí aleatori sinó el conjunt de models ARIMA. Dit d'una altra manera, si es substitueix l'operador *(1-L)* en un model ARIMA per l'operador $(1-L)^d$, s'obté una varietat de models molt més àmplia i flexible que la dels models ARIMA, anomenats models ARFIMA. De fet, els models ARFIMA també es poden interpretar com una ampliació del model SBFI(*d*) per afegir-hi components ARMA de memòria curta.

El punt de partida i referència bàsica pel que fa al desenvolupament dels models ARFIMA és el treball de Hosking (1984) a partir del qual es pot proposar la següent definició.

Definició 4.8. Un procés X_t sense components deterministes es diu que és un ARFIMA(*p*,*d*,*q*) (autoregressiu, integrat fraccionalment i mitjana mòbil) amb $d\in(-0.5,0.5)$ i *p*, *q* naturals, si X_t és una solució estacionària de la següent equació estocàstica en diferències:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \tag{4.47}$$

on ε_t és un procés soroll blanc i $\phi(L)$ i $\theta(L)$ són polinomis en l'operador L amb totes les arrels de l'equació $\phi(z)=0$ i $\theta(z)=0$ fora del cercle unitat.

Sense entrar en detalls, es pot comprovar com un procés d'aquest tipus és estacionari i invertible sempre que $d\in(-0.5,0.5)$. A més, es pot interpretar el paràmetre *d* de manera equivalent a la interpretació en els models SBFI(*d*).

La densitat espectral d'un procés ARFIMA(*p*,*d*,*q*) definit a partir de (4.47) es pot calcular substituint l'operador *L* per $e^{-i\omega}$ de manera que:

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}\left[2sin(\omega/2)\right]^{-2d}\frac{|\theta(\omega)|^2}{|\phi(\omega)|^2} \tag{4.48}$$

Així, s'obté que al voltant de la freqüència zero l'espectre es pot expressar com:

$$f(\omega) \underset{\omega \to 0}{\sim} \frac{\sigma_\varepsilon^2}{2\pi} \frac{|\theta(0)|^2}{|\phi(0)|^2} \omega^{-2d} \tag{4.49}$$

Per tant, si $d \leq 0$ l'espectre està acotat i per tant X_t és de memòria curta mentre que per $d>0$ l'espectre tendeix cap a infinit a la freqüència zero i és de memòria llarga. A diferència dels processos SBFI(d) presentats a l'apartat anterior, quan $d=0$ el procés no queda limitat a soroll blanc, sinó que pot presentar una gran varietat de comportaments ja que pot ser qualsevol model ARMA(p,q).

En tot cas, també té interès generalitzar la definició dels models ARFIMA(p,d,q) a valors de d fora de l'interval en que el procés és estacionari:

Definició 4.9. Un procés X_t sense components deterministes és un ARFIMA(p,d,q) amb $d \geq 0.5$ si existeix un valor d^* natural de manera que $Y_t=(1-L)^{d^*}X_t$ és un procés ARFIMA(p,d-d^*,q) amb $d-d^* \in (-0.5,0.5)$ i, per tant, estacionari.

Per tant, sembla natural modelitzar les sèries integrades d'ordre d amb un model ARFIMA(p,d,q).

S'ha de destacar que es pot generalitzar també pels models ARFIMA(p,d,q) amb $d \geq 0.5$ la noció de pseudo-espectre de manera que l'expressió (4.45) és vàlida per qualsevol ARFIMA.

Tot i que es tracta d'una modelització molt atractiva que, al menys teòricament, permet distingir en una sèrie temporal el comportament associat a la memòria curta del comportament associat a la memòria llarga, un problema important d'aquests models és la dificultat de càlcul d'autocovariàncies, autocorrelacions i autocorrelacions parcials.

En aquest sentit, Hosking (1984) proposa una aproximació molt interessant a aquest problema. Suposi's que X_t és un ARFIMA(p,d,q) $\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$. L'aproximació consisteix a definir dos processos auxiliars Y_t i Z_t en que es permeten separar les components ARMA de la diferència fraccional:

$$(1-L)^d Y_t = \varepsilon_t \tag{4.50}$$

i:

$$\phi(L)Z_t = \theta(L)\varepsilon_t \tag{4.51}$$

Com Z_t es comporta com un procés ARMA, llavors es compleix que:

$$\gamma_k^Z = \phi_1\gamma_k^Z + \phi_2\gamma_{k-1}^Z + \ldots + \phi_p\gamma_{k-p}^Z \tag{4.52}$$

de manera que existeixen dos successions α_j i β_j tals que les autocovariàncies de Z_t es poden calcular per $k>q$ com[13]:

$$\gamma_k^Z = \sum_{j=\Gamma}^{p} \alpha_j \beta_j^{k-q} \tag{4.53}$$

Per la definició de cada un dels processos a partir de transformacions algebraiques s'obté que:

$$\begin{aligned}\gamma_k = \sum_{j=-\infty}^{\infty}\gamma_j^Z\gamma_{k-j}^Y = \ldots = \sum_{j=-q}^{q}\gamma_j^Z\gamma_{k-j}^Y + \gamma_{k+q}^Y\sum_{l=1}^{p}\alpha_l\left[F(1,d-q-k;1-d-q-k;\beta_l)-1\right] + \\ + \gamma_{k-q}^Y\sum_{l=1}^{p}\alpha_l\left[F(1,k-q+d;k-q+1-d;\beta_l)-1\right]\end{aligned} \tag{4.54}$$

on $F(a,b;c;x)$ és la funció hipergeomètrica definida a (4.19).

A partir de (4.54) es pot comprovar com les expressions paramètriques de les representacions autoregressives i mitjana mòbil infinites dels models ARFIMA(p,d,q) són funcions força complicades de la funció hipergeomètrica.

[13] Sobre aquesta propietat dels models ARMA es pot consultar Box i Jenkins (1970). Els coeficients α_j i β_j formen duts successions de coeficients dels que sols ens interessa la seva existència, no la forma com es poden calcular.

A partir de Cheung (1994) es pot trobar una expressió més compacta d'aquestes autocovariàncies basant-se en les autocovariàncies d'un model ARFIMA(0,d,0). Concretament, en aquest treball es demostra que si el procés X_t compleix $\phi(L)(1-L)^d X_t=\theta(L)\varepsilon_t$ les seves autocovariàncies es poden calcular com:[14]

$$\begin{aligned}\gamma_k = \gamma_k^* \sum_{j=1}^{p} \zeta_j \sum_{i=-q}^{q} \psi_i \frac{\Gamma(1-d-k)\Gamma(d-k+p+i)}{\Gamma(d-k)\Gamma(1-d-k+p+i)} \times \\ \times\left[\rho_j^{2p} C(p+i-k,\rho_j) + C(k-p-i,\rho_j) - 1\right]\end{aligned} \tag{4.55}$$

on γ_k^* són les autocovariàncies d'un procés ARFIMA(0,d,0), els coeficients ρ_j són les arrels del polinomi $\phi(L)$ i els coeficient $C(\cdot,\cdot)$ són:

$$C(h,\rho_j) = \sum_{k=0}^{\infty} \frac{\Gamma(1-d+h)\Gamma(d+h+k)}{\Gamma(d+h)\Gamma(1-d+h+k)} \rho_j^k \tag{4.56}$$

Pel que fa als coeficients ζ_j, es poden calcular a partir de:

$$\zeta_j = \left[\rho_j \prod_{i=1}^{p} (1-\rho_i\rho_j) \prod_{m\neq j}^{p} (\rho_j - \rho_m)\right]^{-1} \tag{4.57}$$

i els coeficients ψ_i són:

$$\psi_i = \sum_{k=0}^{q-i} \theta_k \theta_{k+i} \tag{4.58}$$

Observi's que, tot i tractar-se d'una expressió compacte, el seu càlcul és complexe, sobretot degut a la necessitat de calcular els coeficients $C(h,\rho_j)$.[15]

Aquesta complexitat de les autocovariàncies complica l'anàlisi de les característiques dels models ARFIMA i la seva comparació amb les característiques dels models ARMA que és un dels objectius d'aquest apartat.

[14] En el mateix treball es pot consultar també com calcular a la pràctica aquestes autocovariàncies.

[15] De totes maneres, a l'apartat 6.11 de la tesi es proposa un mètode molt simple pel càlcul a nivell pràctic de les autocovariàncies de models ARFIMA.

Per efectuar aquesta comparació, s'analitzen a continuació, i de manera més detallada dos exemples de models ARFIMA molt senzills. Es tracta concretament dels models ARFIMA(1,d,0) i ARFIMA(0,d,1) que, tot i que sols depenen de dos paràmetres, permeten treure conclusions interessants sobre les propietats dels models ARFIMA i mostrar-ne les diferències amb els models ARMA.

A) Model ARFIMA(1,d,0)

Si X_t és un model ARFIMA(1,d,0), llavors $(1-L)^d(1-\phi L)X_t=\varepsilon_t$. Donat que els models SBFI(d) tenen una representació autorregressiva infinita, és natural que els ARFIMA(1,d,0) també. En aquest cas, els coeficients de l'expansió $X_t = \varepsilon_t + \sum \pi_k X_{t-k}$ es pot calcular aprofitant que es compleix que $\pi(L)=(1-L)(1-\phi L)$ i, per tant:[16]

$$\pi_k = \frac{(k-d-2)!}{(k-1)!(-d-1)!}\left\{1-\phi-\frac{1+d}{k}\right\} \tag{4.59}$$

Anàlogament, si hom vol calcular els coeficients de la representació mitjana mòbil infinita del model ARFIMA(1,d,0), com $\psi(L)=\pi(L)^{-1}$, s'obté:

$$\psi_k = \frac{(k+d-1)!}{k!(d-1)!}F(1,-k;1-d-k;\phi) \tag{4.60}$$

Observi's que asimptòticament es compleixen les següents aproximacions:

$$\pi_k \underset{k\to\infty}{\sim} \frac{(1-\phi)}{(-d-1)!}k^{-d-1} \tag{4.61}$$

i:

$$\psi_k \underset{k\to\infty}{\sim} \frac{k^{d-1}}{(-d-1)!(1-\phi)} \tag{4.62}$$

[16] Alguns dels següents resultats es poden consultar, de manera més detallada a Hosking (1981).

de manera que els coeficients π_k i ψ_k tendeixen cap a zero a mesura que k tendeix cap a infinit. De totes maneres, no existeix una bona aproximació a aquesta estructura a partir de models ARMA(p,q) amb p i q finits. Això es pot comprovar millor comparant les autocovariàncies del procés ARFIMA(1,d,0) amb les dels processos ARMA(p,q).

Concretament, si es simplifica l'expressió (4.54) quan $p=1$ i $q=0$, les autocovariàncies d'X_t són iguals a:

$$\gamma_k = \frac{\gamma_k^Y}{(1-\phi^2)}\left[F(1,d+k;1-d+k;\phi)+F(1,d-k;1-d-k;\phi)-1\right] \tag{4.63}$$

on γ_k^Y són les autocovariàncies d'un procés SBFI(d):

$$\gamma_k^Y = \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)}\sigma_\varepsilon^2 \tag{4.64}$$

Per tant:

$$\begin{aligned}\gamma_k = {} & \frac{\Gamma(k+d)\Gamma(1-2d)\sigma_\varepsilon^2}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)(1-\phi^2)}\times \\ & \times\left[F(1,d+k;1-d+k;\phi)+F(1,d-k;1-d-k;\phi)-1\right]\end{aligned} \tag{4.65}$$

A partir d'aquestes autocovariàncies es poden calcular les autocorrelacions:

$$\begin{aligned}\rho_k = {} & \frac{F(1,d+k:1-d+k;\phi)+F(1,d-k;1-d-k;\phi)-1}{(1-\phi)F(1,1+d;1-d;\phi)}\times \\ & \times\frac{d(1+d)(2+d)...(k-1+d)}{(1-d)(2-d)....(k-d)}\end{aligned} \tag{4.66}$$

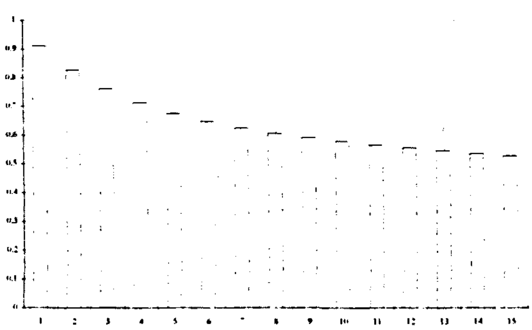
Pel que fa a la densitat espectral de X_t és:

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi(1+\phi^2-2\phi\cos\omega)\left[2\sin(\omega/2)\right]^2} \tag{4.67}$$

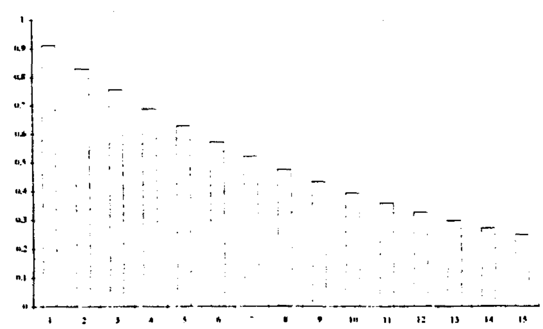
Un aspecte important relacionat amb el models ARFIMA és en quina mesura permeten modelitzar comportaments que no es poden modelitzar amb altres models més senzills. Un bon exemple de la major flexibilitat dels models ARFIMA n'és precisament el model ARFIMA(1,d,0) ja que depèn solament de dos paràmetres.

Així, per exemple, si una sèrie és solució de l'equació $(1-L)^d(1-\phi L)X_t=\varepsilon_t$ amb $\phi=0.5$ i $d=0.4$, el primer coeficient d'autocorrelació és $\rho(1)=0.911$. Als gràfics 4.1 i 4.2 es compara la funció d'autocorrelació simple d'aquest model amb la d'un model AR(1) amb $\phi=0.763$. Observi's que tot i presentar un primer coeficient d'autocorrelació igual els dos gràfics són ben diferents.

Gràfic 4.1. FAS de $(1-L)^{0.4}(1-0.5L)X_t=\varepsilon_t$



Gràfic 4.2. FAS de $(1-0.911L)X_t=\varepsilon_t$



B) Model ARFIMA(0,d,1)

Si X_t és un model ARFIMA(0,d,1), llavors $(1-L)X_t=(1-\theta L)\varepsilon_t$. Noti's que es tracta del model invers d'un model ARFIMA(1,-d,0) amb $(1-\phi L)=(1-\theta L)^{-1}$. Això permet aprofitar els resultats de l'apartat anterior per calcular els coeficients de les representacions autorregressiva, $X_t = \varepsilon_t + \sum \pi_k X_{t-k}$ i mitjana mòbil, $X_t = \sum \psi_k \varepsilon_t - k$. Concretament:

$$\pi_k = \frac{(k-d-1)!}{k!(-d-1)!} F(1,-k;1+d-k;\theta) \tag{4.68}$$

i:

$$\psi_k = \frac{(k+d-2)!}{(k-1)!(d-1)!}\left\{1-\theta-\frac{1-d}{k}\right\} \tag{4.69}$$

Observi's que asimptòticament es compleixen les següents aproximacions:

$$\pi_k \underset{k\to\infty}{\sim} \frac{k^{-d-1}}{(d-1)!(1-\theta)} \tag{4.70}$$

i:

$$\psi_k \underset{k\to\infty}{\sim} \frac{(1-\theta)}{(d-1)!} k^{d-1} \tag{4.71}$$

de manera que els coeficients π_k i ψ_k també tendeixen cap a zero a mesura que k tendeix cap a infinit.

Pel que fa a les autocovariàncies de X_t són iguals a:

$$\gamma_k = \frac{\Gamma(k+d)\Gamma(1-2d)\sigma_\varepsilon^2}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)} \times \frac{(1-\theta)^2 k^2 - (1-d)\left\{(1-d)(1+\theta^2) - 2\theta d\right\}}{k^2 - (1-d)^2} \tag{4.72}$$

i les autocorrelacions:

$$\rho_k = \frac{(1-\theta)^2 k^2 - (1-d)^2 \left\{1+\theta^2 - \dfrac{2\theta d}{1-d}\right\}}{\left[1+\theta^2 - \dfrac{2\theta d}{1-d}\right]\left[k^2 - (1-d)^2\right]} \times \frac{d(1+d)(2+d)\ldots(k-1+d)}{(1-d)(2-d)\ldots.(k-d)} \tag{4.73}$$

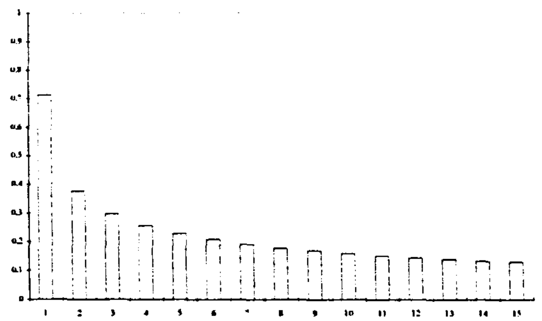
Pel que fa a la densitat espectral de X_t és:

$$f(\omega) = \frac{\sigma_\varepsilon^2(1+\theta^2 - 2\theta\cos\omega)}{2\pi\left[2\sin(\omega/2)\right]^2} \tag{4.74}$$

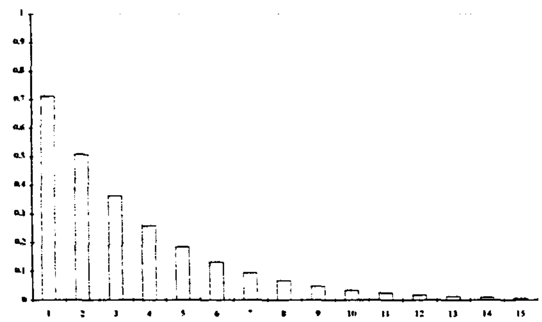
De manera similar al que s'ha fet amb el procés ARFIMA(1,d,0), també té interès comparar les autocorrelacions d'un procés ARFIMA(0,d,1) amb les d'un model AR(1). Així, consideri's a títol d'exemple una sèrie generada a partir de l'equació $(1-L)^d X_t=(1-\theta L)\varepsilon_t$ amb $\theta=-0.9$ i $d=0.25$.

El primer coeficient d'autocorrelació en aquest cas és $\rho(1)=0.713$ i als gràfics 4.3 i 4.4 es compara la funció d'autocorrelació simple d'aquest model amb la d'un model AR(1) amb $\phi=0.713$. Novament, es pot observi's en aquests gràfic com, tot i presentar un primer coeficient d'autocorrelació igual els dos processos presenten característiques molt diferents.

Gràfic 4.3. FAS de $(1\text{-}L)^{0.25}X_t=(1+0.9L)\varepsilon_t$



Gràfic 4.4. FAS de $(1\text{-}0.713L)X_t=\varepsilon_t$



Els resultats anteriors permeten treure les següents conclusions:

- Els models ARFIMA són molt flexibles ja que permeten, amb un nombre reduït de paràmetres, modelitzar sèries amb diferents tipus de memòria, en particular sèries amb memòria llarga que són aquelles per a les que, a nivell teòric, no poden usar-se models ARMA.
- S'ha comprovat com aquests models presenten unes autocorrelacions que no poden aproximarse a través de models ARMA
- Per tant, si hom pretén, tal com és l'objectiu de la tesi, analitzar les conseqüències de tipus economètric que pot tenir la no consideració del tipus de memòria de les dades utilitzades, els models ARFIMA són molt adequats com a eina per aquest tipus d'anàlisi perquè poden assimilar-se els errors en l'especificació del tipus de memòria com errors en la selecció del paràmetre *d* quan les dades s'han generat a través d'un model ARFIMA.

Per les raons anteriors, s'ha optat per utilitzar al llarg de la tesi el supòsit que les dades amb memòria llarga s'han generat a partir d'un model ARFIMA, tot i que la simplificació que això suposa. En tot cas, no sembla que aquest supòsit sigui més restrictiu que el fet de suposar que les dades amb memòria curta es poden considerar com generades per un model ARMA, supòsit habitual a la literatura economètrica.

Tot i la utilització bàsicament dels models ARFIMA, als apartats següents es presenten alguns tipus de models que tot i que també contenen memòria llarga, són molt menys coneguts. De fet, aquests models sols apareixeran de manera molt puntual al llarg de la tesi, però són importants pels següents aspectes:

- Els models ARFIMA no són adequats per la modelització de sèries de freqüència superior a l'anual i que presenten un comportament estacional. Aquest és un aspecte molt important ja que moltes de les variables econòmiques presenten aquesta estacionalitat, però aquest és un aspecte que no serà objecte d'especial consideració al llarg de la tesi (excepte a l'apartat 8.2) ja que no modifica les conseqüències de tipus economètric analitzades a la tesi. Per a la modelització de sèries estacionals amb memòria llarga es poden usar els models GARMA i SARFIMA, presentats breument als apartats 4.6 i 4.7.

- Per raons que s'analitzen en capítols posteriors, els models ARFIMA presenten alguns problemes importants pel que fa a la identificació del tipus de memòria i la realització d'inferència sobre l'ordre d'integració, mentre que els models EXPFI, presentats a l'apartat 4.8 semblen una aproximació interessant a aquest problema que encara no ha estat analitzada i sobre la que s'aporten alguns resultats interessants.

4.6. Els models GARMA

Tal com s'ha analitzat prèviament, els models ARFIMA permeten recollir la presència de memòria llarga en la freqüència zero. Però, en sèries econòmiques amb forta component estacional o amb comportament cíclic molt persistent, és possible la presència de memòria llarga en altres freqüències. Així, Hosking (1981) suggereix, sense formalitzar-ho, estendre els models ARFIMA a altres freqüències diferents de la freqüència zero, considerant:

$$\phi(L)(1-2\xi L+L^2)^\lambda X_t=\theta(L)\varepsilon_t \tag{4.75}$$

L'operador $(1\text{-}2\xi L\text{+}L^2)^\lambda$ es pot escriure en funció dels polinomis de Gegenbauer, molt populars en certes àrees de la matemàtica aplicada.

Definició 4.10. Sigui $\lambda \neq 0$ i $|x|<1$, llavors per qualsevol $|\xi| \leq 1$ es defineixen els polinomis de Gegenbauer $C_k^{(\lambda)}(\xi)$ com:

$$(1-2\xi x+x^2)^{-\lambda}=\sum_{k=0}^{\infty} C_k^{(\lambda)}(\xi)x^k \tag{4.76}$$

Aquests polinomis es poden calcular explícitament. Rainville (1960) mostra que:

$$C_k^{(\lambda)}(\xi)=\sum_{j=0}^{[k/2]}\frac{(-1)^j\Gamma(\lambda+k)(2\xi)^{k-2j}}{j!(k-2j)!\Gamma(\lambda)} \tag{4.77}$$

Definició 4.11. Un procés X_t sense components deterministes, és un procés de Gegenbauer amb paràmetres λ i ξ si:

$$(1-2\xi L+L^2)^{\lambda}X_t=\varepsilon_t \tag{4.78}$$

on ε_t és soroll blanc i λ és diferent de zero.

Noti's que quan $\xi=1$ aquest procés no és més que un soroll blanc fraccional amb $d=2\lambda$.

A Gray et al. (1989) es demostren els següents resultats:

- el procés X_t és estacionari de segon ordre si $|\xi|<1$ i $0<\lambda<0.5$ o si $|\xi|=1$ i $\lambda<0.25$
- el procés X_t és invertible si $|\xi|<1$ i $\lambda>-0.5$ o si $|\xi|=1$ i $\lambda>-0.25$.
- Quan $\xi=1$ es té un procés FI(2λ), de manera que si $0<\lambda<0.25$ les autocorrelacions són:

$$\rho_k=\frac{\Gamma(1-2\lambda)\Gamma(k+2\lambda)}{\Gamma(2\lambda)\Gamma(k-2\lambda+1)} \tag{4.79}$$

de manera que en el límit, quan $k\to\infty$, llavors $\rho_k \sim k^{4\lambda-1}$.

- Quan $\xi=-1$ i $0<\lambda<0.25$ les autocorrelacions són:

$$\rho_k=(-1)^k\frac{\Gamma(1-2\lambda)\Gamma(k+2\lambda)}{\Gamma(2\lambda)\Gamma(k-2\lambda+1)} \tag{4.80}$$

de manera que en el límit, quan $k\to\infty$, llavors $\rho_k \sim (-1)k^{4\lambda-1}$.

- Quan $|\xi|<1$ i $0<\lambda<0.5$ les autocorrelacions compleixen:

$$\rho_k \sim k^{2\lambda-1}\cos(k\omega_0) \tag{4.81}$$

on $\omega_0 = \cos^{-1}(\xi)$ es coneix amb el nom de freqüència *G*.

És immediat comprovar que la densitat espectral és:

$$f(\omega) = \sigma_\varepsilon^2\left\{4(\cos\omega - \xi)^2\right\}^{-\lambda} \tag{4.82}$$

De les característiques d'aquesta densitat espectral es dedueix que un procés de Gegenbauer conté memòria llarga quan $|\xi|=1$ i $0<\lambda<0.25$ o quan $|\xi|<1$ i $0<\lambda<0.5$.

Una característica important a destacar és la diferència radical de comportament entre les freqüències 0 o π i les altres freqüències.

Definició 4.12. Un procés X_t sense components deterministes, és un procés GARMA amb paràmetres λ, ξ, $\phi(L)$ i $\theta(L)$ si:

$$\phi(L)(1-2\xi L + L^2)^\lambda X_t = \theta(L)\varepsilon_t \tag{4.83}$$

on ε_t és soroll blanc i λ és diferent de zero.

Pel que fa a la densitat espectral és:

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}\left\{4(\cos\omega - \xi)^2\right\}^{-\lambda}\left|\frac{\theta(\omega)}{\phi(\omega)}\right|^2 \tag{4.84}$$

de manera que a mesura que ω tendeix cap a ω_0, la densitat espectral tendeix cap a infinit.

Pel que fa a la funció de resposta a l'impuls s'obté asimptòticament:

$$\psi_k \sim \cos[k\omega_0 + d(\omega_0 - 0.5\pi)]k^{d-1} \tag{4.85}$$

4.7. Els models SARFIMA

Una altra possibilitat d'extensió dels models ARFIMA per la modelització de l'estacionalitat és la introduïda per Porter-Hudak (1990) que utilitza una potència fraccional de la diferència estacional coneguts com models SBFIS(d).

Definició 4.13. Un procés X_t sense components deterministes es diu que és un SBFIS(d), soroll blanc integrat fraccionalment estacional amb $d \in (-0.5, 0.5)$, si X_t és una solució estacionària de la següent equació estocàstica en diferències:

$$(1 - L^s)^d X_t = \varepsilon_t \tag{4.86}$$

on ε_t és un procés soroll blanc.

La densitat espectral d'un procés SBFIS(d) és:

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi\{2[1 - \cos(s\omega)]\}^{2d}} \tag{4.87}$$

Si $d>0$, es tracta d'un espectre que tendeix cap a infinit a cada una de les freqüències estacionals. Si $d<1/2$ és estacionari amb representació mitjana mòbil, $X_t = \sum \psi_k \varepsilon_{t-k}$ amb coeficients no nuls sols per retards múltiples de s:

$$\psi_{sk} = \frac{\Gamma(sk - d)}{\Gamma(sk + 1)\Gamma(-d)} = \sigma^2 \cos(k\pi/s) \frac{\Gamma(1 - d + sk)}{\Gamma(1 - d - k/s)} \tag{4.88}$$

Pel que fa a les autocovariàncies són:

$$\gamma_k = \sigma^2 \cos(k\pi/s) \frac{\Gamma(1 - 2d)}{\Gamma(1 - d + k/s)\Gamma(1 - d - k/s)} \tag{4.89}$$

De manera similar als models presentats abans, també es pot generalitzar aquest model estacional per incorporar estructures de memòria curta a través de models ARMA[17]:

[17] Ray (1993b) aplica els models SARFIMA per a l'estimació de prediccions en sèries econòmiques.

Definició 4.14. Un procés X_t sense components deterministes es diu que és un SARFIMA(p,d,q), autoregressiu, integrat fraccionalment i mitjana mòbil estacional amb ordre d'integració d on $d\in(-0.5,0.5)$, si X_t és una solució estacionària de la següent equació estocàstica en diferències:

$$\phi(L)(1-L^s)^d X_t = \theta(L)\varepsilon_t \tag{4.90}$$

on ε_t és un procés soroll blanc.

La densitat espectral d'aquest procés és:

$$f(\omega)=\frac{\sigma_\varepsilon^2}{2\pi\{2[1-\cos(s\omega)]\}^{2d}}\left|\frac{\theta(\omega)}{\phi(\omega)}\right|^2 \tag{4.91}$$

4.8. Els models EXPFI

Tot i que l'ajust de models paramètrics a les sèries temporals ha generat molta discussió gairebé sempre s'han usat els models ARMA analitzats per Box i Jenkins (1970). En aquest cas, es suposa que la sèrie observada X_t està relacionada amb un soroll blanc ε_t segons la següent equació:

$$(1-\sum_{k=1}^{p}\phi_k L^k)X_t = (1-\sum_{k=1}^{q}\theta_k L^k)\varepsilon_t \tag{4.92}$$

Recordi's que la densitat espectral d'aquestes sèries és de la forma:

$$f(\omega)=\frac{\sigma^2}{2\pi}\left|\frac{1-\sum_{k=1}^{q}\theta_k e^{ik\omega}}{1-\sum_{k=1}^{p}\phi_k e^{ik\omega}}\right|^2 \tag{4.93}$$

Hi ha, però, un altre model menys popular a l'Econometria, proposat inicialment per Bloomfield (1973) que es basa en l'observació empírica de que el logaritme de la densitat espectral estimada, té habitualment un comportament molt regular de manera que es pot aproximar per una sèrie de Fourier truncada. Es tracta d'un model paramètric definit en el domini de les freqüències.

Definició 4.15. Un procés X_t sense components deterministes es diu que és exponencial o EXP(p) si la seva densitat espectral és[18]:

$$f(\omega) = \frac{\sigma^2}{2\pi} \exp\left\{ 2\sum_{k=1}^{p} \beta_k \cos(k\omega) \right\} \tag{4.94}$$

De fet, aquest model no ha estat gens popular entre els econòmetres però com s'ha observat a Gil-Alaña i Robinson (1997) té unes propietats molt interessants per la discriminació entre els diferents tipus de memòria.

De fet, tal com s'analitza en capítols posteriors de la tesi, té força interès usar una extensió d'aquest model que permeti modelitzar tant la presència de memòria llarga com de memòria curta simultàniament.

La forma més senzilla de considerar aquesta generalització és definir el concepte de procés exponencial fraccionalment integrat.

Definició 4.16. Un procés X_t sense components deterministes es diu que és un procés exponencial fraccionalment integrat d'ordre d, EXPFI(p,d), si el procés $Y_t=(1-L)^d X_t$ té densitat espectral igual a:

$$f_Y(\omega) = \frac{\sigma^2}{2\pi} \exp\left\{ 2\sum_{k=1}^{p} \beta_k \cos(k\omega) \right\} \tag{4.95}$$

[18] N'hi ha prou amb usar sols termes amb cosinus perquè es tracta d'una funció parell.

Per tant, la densitat espectral d'un procés EXPFI(p,d) és:

$$f(\omega)=\frac{\sigma^2}{2\pi}\exp\left\{2\sum_{k=1}^{p}\beta_k\cos(k\omega)\right\}\left\{4sin^2(\omega/2)\right\}^{-d} \tag{4.96}$$

A Beran (1993) es proposen un conjunt de models amb memòria llarga amb una definició similar a aquesta. Malauradament, aquests models no s'han usat a l'anàlisi economètrica tot i que als apartats 6.13 i 7.5 es proposen algunes aplicacions d'aquests models que semblen especialment interessants i obren línies de recerca importants.

4.9. Consideracions finals

A partir dels resultats presentats als apartats anteriors, si hom es centra en l'anàlisi de les conseqüències dels diferents tipus de memòria en la freqüència zero, sembla raonable basar-se en aquests models ARFIMA ja que presenten una gran flexibilitat i contenen com a casos particulars sèries amb els diferents tipus de memòria definits al capítol 3.

De totes maneres, tot i que a l'apartat 4.5 de la tesi s'han analitzat algunes de les característiques d'aquests models és necessari destacar-ne alguns aspectes addicionals abans de continuar amb els objectius de la tesi. Així, al quadre 4.1 es recullen les principals característiques dels models ARFIMA(p,d,q) segons quin sigui el valor del paràmetre d, i suposant sempre que totes les arrels dels polinomis autoregressiu i mitjana mòbil es troben fora del cercle unitat.

En aquest quadre apareixen algunes característiques no analitzades fins al moment i que es justifiquen breument a continuació. Així, pel que fa a la funció de resposta a l'impuls del model ARFIMA(p,d,q), permet conèixer si la sèrie presenta o no reversió a la mitjana. Com han mostrat Campbell i Mankiw (1987), si es defineix:

$$(1-L)X_t=A(L)\varepsilon_t \tag{4.97}$$

la possible reversió depèn del terme $A(1)$. Com:

$$A(L)=(1-L)^{1-d}\phi(L)^{-1}\theta(L)=F(d-1,1;1;L)\phi(L)^{-1}\theta(L) \tag{4.98}$$

A partir de Gradszteyn i Ryzhnik (1980 pp. 1039-1042) es comprova que $F(d\text{-}1,1;1;L)=0$ si $d<1$, i per tant $A(1)=0$ sempre que $d<1$. Més concretament, l'efecte d'un *shock* produït en el moment sobre el valor de la variable en el moment $t+k$ és $1+A_1+A_2+...A_k$. Per tant, per a que un procés presenti reversió a la mitjana cal que $A(1)=0$. Així, sempre que $d<1$, el procés presenta reversió a la mitjana.

Quadre 4.1. Característiques principals dels models ARFIMA(p,d,q) segons el valor de d.

	-0.5<d<0	**d=0**	**0<d<0.5**	**0.5≤d<1**	**d=1**	**d>1**
Model	ARFIMA	ARMA	ARFIMA	ARFIMA	ARIMA	ARFIMA
Memòria	curta	curta	llarga	perm. finita	perm. infinita	perm. infinita
Estacionarietat	SI	SI	SI	NO	NO	NO
Invertibilitat	SI	SI	SI	SI	SI	SI
Correlograma	hiperbòlic	exponencial	hiperbòlic	constant	constant	constant
Suma correlacions	finita	finita	infinita	infinita	infinita	infinita
Variància	constant	constant	constant	$\rightarrow\infty$	$\rightarrow\infty$	$\rightarrow\infty$
Efecte shocks	transitoris	transitoris	transitoris	permanents	permanents	permanents
Persistència	*antipersistència*	petita	gran	petita	infinita	infinita
Espectre	finit	finit	infinit	infinit	infinit	infinit
Reversió a la mitjana	SI	SI	SI	SI	NO	NO
Temps esperat per creuar $E(X_t)$	poc	poc	finit	finit	infinit	infinit
Sumes parcials	< lineal	lineal	$O(T^{1+2d})$	$O(T^{1+2d})$	$O(T^3)$	$O(T^{1+2d})$

Una altra propietat força interessant atenent a les mesures presentades al capítol 2, és el comportament de les sumes parcials, $S_T = X_1+...+X_T$. Concretament, a Sowell (1992) es demostra que en un procés ARFIMA(0,d,0), quan $-0.5<d<0.5$:

$$\mathrm{var}(S_T) = \sigma^2 \frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+2d)\Gamma(1+d)} c_T \tag{4.99}$$

on:

$$c_T = \frac{\Gamma(1+d+T)}{\Gamma(T-d)} - \frac{\Gamma(1+d)}{\Gamma(-d)} \tag{4.100}$$

i que asimptòticament es compleix:

$$\lim_{T\to\infty} \frac{\text{var}(S_T)}{T^{1+2d}} = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(1+2d)\Gamma(1+d)\Gamma(1-d)} \tag{4.101}$$

Per tant, quan $-0.5<d<0.5$ l'ordre de creixement de les sumes parcials és[19]:

$$\text{var}(S_T) = O(T^{2d+1}) \tag{4.102}$$

A Diebold i Lindner (1996) es generalitza aquesta propietat per qualsevol model AFIMA(p,d,q) amb $-0.5<d<0.5$. Per tant, en el cas $d=0$, la variància de la suma parcial creix de manera lineal. Per a processos de memòria curta *antipersistent*, quan $-0.5<d<0$, la variància creix a una velocitat inferior, mentre que per a processos de memòria llarga aquest creixement és més ràpid. Observi's, en particular, que quan d s'apropa a $1/2$, la successió S_T s'apropa cap a $O(T^2)$.

Pel cas de models ARFIMA(p,d,q) amb $d\geq 1/2$ pot generalitzar-se fàcilment el resultat anterior i comprovar que les sumes parcials són també $O(T^{2d+1})$. Concretament, si $(1-L)^d X_t=\psi(L)\varepsilon_t$, es pot aprofitar que $X_t=(1-L)Y_t$ on $Y_t\sim$ ARFIMA(p,d^*,q) amb $d^*=d-1$ i, per tant, $-0.5<d^*<0.5$. Així, es poden escriure les sumes parcials S_T com:

$$S_T = \sum_{i=1}^{T} X_i = \sum_{i=1}^{T}\sum_{j=1}^{i-1} Y_{i+j} = \sum_{i=1}^{T}(T-i)Y_i \tag{4.103}$$

de manera que la variància d'aquestes sumes parcials és:

[19] Es diu que una successió $\{a_n\}$ és almenys d'ordre n^λ, notat com $a_n=O(n^\lambda)$, quan existeixen un valor real M amb $0<M<\infty$ i un valor natural N de manera que per $n\geq N$, es compleix que $|n^{-\lambda}a_n|<M$.

$$\begin{aligned}\operatorname{var}(S_T) &= \sum_{i=1}^{T}\sum_{j=1}^{T}(T-i)(T-j)\operatorname{cov}(Y_iY_j) = \sum_{i=1}^{T}(T-i)\left[\sum_{j=1}^{T}(T-j)\gamma^{Y}_{i-j}\right] = \\ &= \sum_{i=1}^{T}(T-i)\left[\sum_{j=1}^{T}(T-j)O(T^{2d^*+1})\right] = \\ &= \sum_{i=1}^{T}(T-i)\left[\sum_{j=1}^{T}(T-j)O(T^{2d^*-1})\right] = \sum_{i=1}^{T}(T-i)O(T^{2d^*}) = O(T^{2d^*+1})\end{aligned} \tag{4.104}$$

Pel que fa a la funció variància-temps[20], $\sigma^2_k = var(X_t\text{-}X_{t-1})$, també presenta un comportament diferent segons quin sigui el valor de d. De fet, es compleix que $\sigma^2_k \sim O(k^{2d-1})$. Per tant, l'estadístic $R(k)$, definit a l'apartat 2.4 de la tesi, també presenta diferències. Així, si $d<1/2$ llavors $R(k)$ és decreixent; si $1/2\leq d<1$, $R(k)$ augmenta a una taxa decreixent; mentre que per $1\leq d<3/2$ augmenta a una taxa creixent.

En relació amb la reversió a la mitjana, tot i que per a qualsevol procés ARFIMA(p,d,q) amb $d<1$ hi ha reversió a la mitjana, aquesta és més o menys ràpida segons el valor de d. És a dir, tot i que el temps esperat per a que la sèrie torni al seu valor mitjà és finit en tots els casos és molt diferent segons si la memòria és curta, llarga o permanent finita.

Com a exemple senzill d'aquestes diferències tot i haver-hi aquesta reversió a la mitjana, s'han generat 10000 sèries segons models ARFIMA(1,0,0) i ARFIMA(0,d,0) amb diferents valors dels paràmetres ϕ i d, totes elles amb $X_0=0=E(X_t)$. Per cada una d'aquestes 10000 sèries s'ha calculat el número d'observacions que han estat necessàries per a que la sèrie creui una línia horitzontal imaginària centrada en el zero, i s'ha calculat el valor mitjà d'aquest número d'observacions. Els resultats pels diferents paràmetres es presenten al quadre 4.2.

Pel que fa al grau de persistència dels shocks en models ARFIMA cal fer un comentari addicional. Tot i que en general, quan $d>0$ els models ARFIMA presenten una persistència major que per models ARMA, cal tenir en compte que la presència de paràmetres autoregressius i mitjana mòbil poden arribar a compensar gairebé l'efecte del paràmetre d.

[20] Vegi's l'apartat 2.4. de la tesi.

Quadre 4.2. Temps mitjà empíric fins creuar el valor esperat.

Model ARFIMA(0,d,0)		Model ARFIMA(1,0,0)	
d	Temps mitjà	ϕ	Temps mitjà
0.00	6.88	0.00	6.78
0.05	7.86	0.10	6.98
0.10	8.33	0.20	7.60
0.15	8.48	0.30	7.64
0.20	8.59	0.40	7.82
0.25	10.40	0.50	8.51
0.30	12.33	0.60	9.19
0.35	17.91	0.70	9.60
0.40	25.45	0.80	9.93
0.45	40.30	0.90	14.93

Per analitzar aquest aspecte és adequat basar-se novament en la representació de Wold:

$$X_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{4.105}$$

A partir de la successió de coeficients ψ_j es poden usar varies mesures representatives del grau de persistència. Les més habituals són les següents:

- Retard mitjà. Es defineix com la mitjana dels diferents retards ponderats per la importància de cada un d'aquests retards en la funció de memòria:

$$\bar{r} = \frac{\sum_{j=0}^{\infty} j\psi_j}{\sum_{j=0}^{\infty} \psi_j} \tag{4.106}$$

- Retard medià. Es defineix com el retard que acumula el 50% de l'efecte total del shock.

Aquestes mesures presenten un problema important en el cas de processos amb memòria llarga o permanent, quan la suma dels coeficients ψ_j és infinita perquè llavors no es poden calcular ni el retard mitjà ni el retard medià.

En aquests casos és especialment interessant mesurar la persistència a partir del comportament dels errors de predicció. En primer lloc, suposant que X_t és estacionari, la variància del soroll blanc i de la sèrie X_t estan relacionades segons:

$$\sigma_\varepsilon^2 = \sigma_X^2 \left[\sum_{j=0}^{\infty} \psi_j^2 \right]^{-1} \tag{4.107}$$

Tal i com s'ha presentat al capítol 3, es pot mesurar la memòria a partir de la variància de l'error de predicció. Quan es vol predir el valor d'X_{T+k} a partir d'observacions $X_1,..,X_T$, aquesta variància és:

$$\xi(k) = \sigma_\varepsilon^2 \sum_{j=0}^{k-1} \psi_j^2 \tag{4.108}$$

Òbviament, a mesura que augmenta la distància temporal k per a la que vol obtenir-se una predicció, tot i que pugui disposar-se d'un predictor sense biaix, la seva variància va augmentant com es pot observar a l'expressió (4.108). Observi's també que el límit de $\xi(k)$ a mesura que k augmenta és la variància del procés X_t. Dit d'una altra manera, si vol predir-se un valor molt llunyà, arriba un moment en que les dades disponibles de la variable no permeten reduir la incertesa sobre aquest valor.

Naturalment, com major sigui la memòria d'una variable, major serà el número d'observacions futures sobre el que es poden calcular prediccions ja que major és la màxima distància temporal a través de la qual es manté una certa relació entre els valors de la variable. En aquest context si hom està interessat en una determinada variable X_t, té interès preguntar-se quin és el màxim valor de k per al que té sentit calcular prediccions, entenent que sols té sentit quan la variància de la predicció sigui inferior a la variància de la sèrie.

A partir d'aquesta idea, es pot definir el concepte de *predicció màxima* com aquell valor M que compleix:

$$\xi(M) < (1-\alpha)\sigma_X^2 < \xi(M+1) \tag{4.109}$$

de manera que M és el nombre màxim d'observacions que es poden predir amb una variància inferior al 100 (1-α)% de la variància de la sèrie. Òbviament, per calcular M cal fixar α com un valor arbitrari tal que $\alpha\in(0,1)$ però un cop fixat, permet comparar la predicció màxima entre diferents models. Així, per exemple si s'usa un valor $\alpha=0.05$, el valor M és el nombre de prediccions que poden calcular-se sense que la seva variància superi el 95% de la variància de la sèrie.

A continuació, s'utilitza aquesta mesura per comparar el grau de persistència que presenten els diferents models ARFIMA segons el valor dels seus paràmetres. Per fer aquesta comparació, s'ha fixat un valor $\alpha=0.05$, i s'ha calculat l'estadístic M a partir del model ARFIMA(1,d,1):

$$(1-L)^d(1-\phi L)X_t = (1-\theta L)\varepsilon_t \tag{4.110}$$

per diferents valors dels paràmetres ϕ, θ i d.

Dels quadres 4.3 a 4.6, on es presenten alguns dels resultats obtinguts se'n poden treure conclusions molt interessants:

- Quan $d=0$, és a dir, quan les dades s'han generat com un model ARMA(1,1), tot i la presència d'un coeficient autoregressiu la variància de les prediccions tendeix ràpidament a la variància de la sèrie. En aquest cas és el paràmetre autoregressiu el que determina la memòria de la sèrie i no el paràmetre mitjana mòbil de manera que per valors de ϕ inferiors a 0.8 en valor absolut sols té sentit predir com a molt 2 o 3 observacions.

- En canvi, el paràmetre d introdueix una major memòria. De fet, quan les dades s'han generat com un model ARFIMA(0,d,0), la predicció màxima passa de $M=0$ quan $d=0$ a $M=37$ quan $d=0.3$.

- De manera sorprenent, la presència d'un paràmetre autoregressiu negatiu i gran en valor absolut elimina part de la memòria introduïda per la diferència fraccional. Per exemple, en un ARFIMA(1,d,0), la màxima predicció per $d=0.3$ i $\phi=-0.9$ queda reduïda a $M=14$.

Quadre 4.3. Predicció màxima *M* per models ARFIMA(1,*d*,1) amb *d=0*.

$\phi \backslash \theta$	-0.90	-0.80	-0.70	-0.60	-0.50	-0.40	-0.30	-0.20	-0.10	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
-0.90	14	14	14	14	14	14	14	14	14	14	13	13	13	12	11	9	6	1	0
-0.80	7	7	7	7	7	7	7	7	6	6	6	6	5	5	4	2	0	0	0
-0.70	4	4	4	4	4	4	4	4	4	4	3	3	3	2	1	0	0	0	1
-0.60	3	3	3	3	3	3	3	3	3	2	2	2	1	1	0	0	0	1	1
-0.50	2	2	2	2	2	2	2	2	2	2	1	1	1	0	0	0	1	1	1
-0.40	2	2	2	2	2	2	2	1	1	1	1	0	0	0	0	0	1	1	1
-0.30	2	2	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1
-0.20	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
-0.10	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1
0.00	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1
0.10	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1
0.20	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1
0.30	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	2	2
0.40	1	1	1	0	0	0	0	0	1	1	1	1	2	2	2	2	2	2	2
0.50	1	1	1	0	0	0	1	1	1	2	2	2	2	2	2	2	2	2	2
0.60	1	1	0	0	0	1	1	2	2	2	3	3	3	3	3	3	3	3	3
0.70	1	0	0	0	1	2	3	3	3	4	4	4	4	4	4	4	4	4	4
0.80	0	0	0	2	4	5	5	6	6	6	6	7	7	7	7	7	7	7	7
0.90	0	1	6	9	11	12	13	13	13	14	14	14	14	14	14	14	14	14	14

Quadre 4.4. Predicció màxima *M* per models ARFIMA(1,*d*,1) amb *d=0.1*.

$\phi \backslash \theta$	-0.90	-0.80	-0.70	-0.60	-0.50	-0.40	-0.30	-0.20	-0.10	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
-0.90	14	14	14	14	14	14	14	14	14	14	14	14	13	12	11	9	6	0	0
-0.80	7	7	7	7	7	7	7	7	6	6	6	6	5	4	4	2	0	0	1
-0.70	4	4	4	4	4	4	4	4	4	4	3	3	2	2	1	0	0	1	1
-0.60	3	3	3	3	3	3	3	3	2	2	2	2	2	0	0	0	0	1	1
-0.50	2	2	2	2	2	2	2	2	2	2	1	1	0	0	0	0	1	1	1
-0.40	2	2	2	2	2	2	2	1	1	1	1	0	0	0	0	1	1	1	1
-0.30	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1
-0.20	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1
-0.10	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1
0.00	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1
0.10	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1
0.20	1	1	1	1	0	0	0	0	1	1	1	1	1	2	2	2	2	2	2
0.30	1	1	1	0	0	0	0	1	1	1	2	2	2	2	2	2	2	2	2
0.40	1	1	0	0	0	0	1	1	2	2	2	2	2	3	3	3	3	3	3
0.50	1	1	0	0	0	1	2	2	2	3	3	3	3	3	3	3	4	4	4
0.60	1	0	0	0	1	2	3	3	4	4	4	4	4	4	5	5	5	5	5
0.70	0	0	0	1	3	4	5	5	6	6	6	6	6	6	6	6	7	7	7
0.80	0	0	2	5	6	7	8	9	9	9	9	9	10	10	10	10	10	10	10
0.90	0	5	11	14	18	19	19	20	21	21	21	21	21	22	22	22	22	22	22

Quadre 4.5. Predicció màxima *M* per models ARFIMA(1,*d*,1) amb *d=0.2.*

$\phi \backslash \theta$	-0.90	-0.80	-0.70	-0.60	-0.50	-0.40	-0.30	-0.20	-0.10	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
-0.90	19	18	18	18	18	18	18	18	18	17	17	16	16	15	13	11	8	4	2
-0.80	7	7	7	7	7	7	7	7	7	6	6	6	6	5	4	4	2	2	3
-0.70	4	4	4	4	4	4	4	4	4	4	4	4	3	2	2	2	2	3	3
-0.60	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	3	3
-0.50	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3
-0.40	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3
-0.30	1	1	1	1	1	1	1	1	1	1	0	2	2	2	3	3	3	3	3
-0.20	1	1	1	1	1	1	1	1	0	0	1	2	2	3	3	3	3	3	3
-0.10	1	1	1	1	1	1	1	0	0	1	2	2	3	3	3	4	4	4	4
0.00	1	1	1	1	1	1	0	0	1	2	2	3	3	4	4	4	4	4	4
0.10	1	1	1	1	0	0	0	0	2	2	3	3	4	4	5	5	5	5	5
0.20	1	1	1	0	0	0	0	2	3	3	4	4	5	5	5	6	6	6	6
0.30	1	1	0	0	0	0	2	3	4	4	5	5	6	6	6	7	7	7	7
0.40	1	1	0	0	0	2	3	4	5	6	6	7	7	8	8	8	8	8	8
0.50	1	0	0	0	2	4	5	6	7	8	8	9	9	10	10	10	10	10	10
0.60	0	0	0	2	4	6	8	9	10	11	11	12	12	12	13	13	13	13	13
0.70	0	0	2	5	8	10	12	13	14	15	15	15	16	16	16	16	16	16	16
0.80	0	2	8	12	14	15	16	17	18	18	18	18	18	19	19	19	19	19	19
0.90	2	13	17	19	20	25	29	37	56	56	57	57	57	57	58	58	58	58	58

Quadre 4.6. Predicció màxima *M* per models ARFIMA(1,*d*,1) amb *d=0.3.*

$\phi \backslash \theta$	-0.90	-0.80	-0.70	-0.60	-0.50	-0.40	-0.30	-0.20	-0.10	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
-0.90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	58	95	49	37
-0.80	9	9	9	9	9	9	10	10	10	10	12	12	16	22	27	31	35	37	38
-0.70	5	4	4	4	4	4	4	4	4	6	6	8	12	18	25	32	37	41	43
-0.60	3	3	3	3	3	3	3	3	4	4	6	11	17	23	31	37	42	46	48
-0.50	2	2	2	2	2	2	2	2	4	6	10	16	23	30	37	43	48	51	52
-0.40	2	2	2	2	2	2	2	2	5	9	15	22	29	37	44	49	53	56	57
-0.30	1	1	1	1	1	2	2	4	8	14	21	29	37	44	51	56	59	61	63
-0.20	1	1	1	1	1	2	3	7	13	20	28	37	45	52	58	63	66	68	69
-0.10	1	1	1	1	1	2	5	11	18	27	37	46	55	62	67	71	74	76	77
0.00	1	1	1	1	1	4	9	17	26	37	47	57	65	72	77	81	84	85	86
0.10	1	1	1	1	1	7	15	25	37	49	60	70	78	85	90	93	95	97	98
0.20	1	1	1	0	5	13	24	37	51	65	76	86	94	100	105	108	110	111	112
0.30	1	1	0	0	10	22	37	54	70	85	97	107	114	120	124	127	129	130	131
0.40	1	0	0	7	19	37	57	78	96	112	124	134	141	146	150	152	154	155	156
0.50	0	0	1	16	37	63	89	113	133	149	161	170	176	181	184	186	188	189	189
0.60	0	0	12	37	71	105	136	161	180	194	204	211	217	221	223	225	226	227	227
0.70	0	4	37	96	179	231	245	278	322	323	324	325	325	325	326	327	327	328	328
0.80	1	37	133	226	453	478	513	536	576	625	636	642	647	649	650	651	652	653	653
0.90	37	184	323	631	905	963	998	1035	1041	1051	1057	1061	1064	1065	1066	1067	1068	1069	1069

- Però encara més sorprenent és que la presència d'un coeficient mitjana mòbil positiu també absorbeix bona part de la memòria introduïda per la diferència fraccional.

- Finalment, de la combinació d'un paràmetre autoregressiu positiu superior a 0.5 i d'un paràmetre d'integració fraccional a partir de $d=0.3$, sorgeixen processos estocàstics amb una predicció màxima superior a 200 i fins i tot 1000 observacions.

Per tant, tot i la utilitat de la classificació presentada de les sèries temporals atenent al tipus de memòria, cal anar amb cura amb la interpretació d'aquesta classificació. Així, dintre de cada categoria hom es pot trobar amb una gran diversitat pel que fa a la quantitat de memòria que presenten les sèries. En tot cas, és de preveure que aquesta diversitat de situacions pugui ser molt més complexa a mesura que augmenta el nombre de paràmetres usats. Per tant, és important també aplicar en la modelització ARFIMA el principi de parsimònia.

Aquests resultats anuncien també que el problema de la identificació en el cas de models ARFIMA promet ser una tasca força complexa atès que la presència de paràmetres autoregressius o mitjana mòbil pot tendir a compensar la presència del paràmetre d. Així, es dedica el capítols següent a la recerca d'algun mètode que permeti identificar el tipus de memòria de les variables econòmiques.