UNIVERSITAT DE BARCELONA

Facultat de Ciències Econòmiques i Empresarials
Departament de Matemàtica Econòmica, Financera i Actuarial

LA MEMÒRIA ALS MERCATS FINANCERS: UNA ANÀLISI MITJANÇANT XARXES NEURALS ARTIFICIALS

Mª Teresa Sorrosal i Forradellas

2005

Capítol 3. LA MEMÒRIA-DEPENDÈNCIA

"A second problem with the econometric view of the world is its treatment of time. It ignores time or, at best, treats time as a variable like any other. The markets and the economy have no memory, or only limited memory, of the past."

Edgar E. Peters (1991:5)

3.1. Introducció.

Al capítol segon resumíem el significat i distingíem entre diferents tipus de memòria dins un context molt ampli, que abastava principalment la vessant biològica. A continuació ens centrem ja en el tractament atorgat a la memòria dins l'entorn financer, començant per la primera accepció destacada, la memòria-dependència.

Des de que sorgeix la necessitat de considerar la influència de les dades del passat en l'evolució futura d'una variable, bàsicament com a crítica a la hipòtesi d'independència, s'ha desenvolupat una àmplia tasca de recerca sobre la memòria, entesa com a una propietat particular d'una sèrie temporal, dins el camp de l'economia i les finances. L'Econometria ha estat la disciplina encarregada de l'elaboració de models cada cop més complexos per a ajustar el comportament real dels preus d'un actiu financer a un model teòric que incorpori els efectes dels preus passats del mateix actiu.

L'objectiu del capítol és analitzar la complexitat de la memòria-dependència i detectar possibles relacions amb la memòria col·lectiva mitjançant l'ús de les XNA. Així, en primer lloc, recollim l'origen i contextualització dels estudis sobre la influència d'observacions passades en una sèrie temporal. La formalització matemàtica de la memòria-dependència és problemàtica, com així ho demostra la varietat de definicions que s'estableixen en funció de tants altres criteris, alguns dels quals s'exposen posteriorment a mode d'exemple. Després, farem un repàs dels models més usuals per a ajustar sèries que presentin aquesta propietat i comentarem els resultats que s'han obtingut a partir d'ells en l'àmbit financer.

Finalment, mostrarem una aplicació en la que hem utilitzat les xarxes *back-propagation* amb un doble fi: presentar un instrument amb el que també és possible modelitzar conjunts de dades ordenades cronològicament i, d'altra banda, emprar aquesta mateixa eina per a analitzar les conseqüències de la inclusió d'elements de la memòria col·lectiva sobre els efectes de la memòria-dependència[1]. En concret, hem examinat la influència de l'"efecte octubre" en l'evolució del *Dow Jones Industrial Average*.

3.2. Antecedents.

3.2.1. La Hipòtesi dels Mercats Eficients.

La presència de la *hipòtesi dels mercats eficients* (HME) en la nostra investigació és deguda a un motiu prou important: representa el punt de partida i pol oposat a la memòria que s'estudia en les sèries temporals[2]. Segons aquesta hipòtesi, el preu actual d'un actiu recull completament tota la informació disponible sobre ell fins aquest moment. Parlant en termes estadístics diríem que el preu és, en cada instant, el millor estimador del valor intrínsec de l'actiu considerat.

La HME es presenta sota tres formes:

- en *sentit dèbil*: el conjunt de tota la informació disponible es limita únicament a la sèrie temporal que recull els preus o les rendibilitats històriques de l'actiu;

[1] La memoria-patró, per la seva naturalesa, no pot participar en la comparació amb la memòria-dependència mitjançant les variacions produïdes en les matrius de pesos d'una xarxa *back-propagation*. Les relacions entre ambdues accepcions són analitzades en el capítol quart.

[2] Encara que la idea pionera de que les variacions en els preus dels actius financers segueixen un camí aleatori fou ja plantejada per Louis Bachelier l'any 1900, i a pesar dels treballs empírics i teòrics anteriors d'investigadors com Kendall, Working, Roberts o Osborne, fou a partir dels anys 60 amb els treballs de Paul A. Samuelson i Benoit Mandelbrot, i sobretot al 1970 de la mà d'Eugene Fama quan es donà a conèixer la HME. Encara avui dia forma la base de models financers àmpliament utilitzats per agents econòmics que actuen diàriament als parquets, com són el *Capital Asset Pricing Model*, l'*Arbitrage Pricing Theory* o el model de valoració d'opcions de Black-Scholes.

- en *sentit semi fort*: els preus també s'ajusten a qualsevol altra informació disponible de caire públic, és a dir, a l'abast de tots els agents econòmics que participen en el mercat;
- en *sentit fort*: a la interpretació anterior s'afegeix la resta d'informació amb rellevància sobre l'actiu de referència, encara que aquesta estigui en mans d'una minoria d'individus del mercat, la informació privilegiada.

En qualsevol dels tres sentits és necessari que tota informació rellevant que arribi al mercat immediatament repercuteixi en el preu de l'actiu, de manera que s'evitin possibilitats d'arbitratge[3] i el mercat romangui en equilibri. Per altra banda, es suposa que els inversors són individus “racionals”[4] i com a tals reaccionaran interpretant correctament les noves notícies, anant les seves decisions encaminades a corregir el diferencial creat entre preu i valor de l'actiu, interpretat aquest darrer com la suma dels futurs fluxos monetaris generats per l'actiu i descomptats al moment actual.

Sobre la racionalitat dels agents econòmics se n'han fet crítiques molt suggeridores. Cal destacar les d'Edgar E. Peters (1999:78-81) i Daniel McFadden (1999) per la seva relació amb el tema que aquí ens ocupa. Entre els aspectes més discutibles per acceptar que els individus es comporten racionalment citen l'existència de retards temporals entre l'arribada de nova informació al mercat i la reacció dels agents, la valoració esbiaixada i subjectiva del contingut de la informació, la heterogeneïtat en el comportament dels agents davant el risc i la influència de la pertinença a un col·lectiu en la presa de decisions.

El supòsit subjacent de la HME al que prestarem més atenció no és el de la racionalitat dels individus, sinó el d'independència entre els increments relatius en el preu dels

[3] Una de les crítiques que s'han fet a la hipòtesi dels mercats eficients és que la constant adequació del preu de l'actiu financer al seu valor intrínsec anul·laria completament qualsevol possibilitat de negoci, i per tant, el volum de negociació seria zero. E. Fama respon que és precisament aquest fet, que els agents creguin que els mercats són ineficients, el que fa que en conjunt siguin eficients.
D'altra banda, l'existència de memòria-dependència no implica l'aparició de possibilitats d'arbitratge. Poden donar-se en models com el moviment brownià fraccional, però altres propostes eliminen aquesta condició. Vegi's Rogers, L.C.G. (1997).
[4] El terme 'racionalitat' ha estat posat en entredit moltes vegades al llarg de la història de l'economia, degut a la problemàtica sobre què s'ha de considerar un “comportament racional". En Allais, M. (1953) i en McFadden, D. (1999) trobem dos bons exemples pertanyents a diferents èpoques però versant sobre la mateixa matèria. Nosaltres utilitzem aquí "racionalitat" en el mateix sentit que és emprat en la HME sense entrar en altres consideracions. Així, l'entenem com a forma d'actuar dirigida a la convergència entre preu de mercat i el valor intrínsec de l'actiu, tot cercant la maximització de la rendibilitat esperada donat un nivell de risc.

actius financers; amb el benentès de que implícitament els elements que intervenen en el primer són també factors determinants del segon[5].

3.2.2. La hipòtesi d'independència i el model de passeig aleatori.

La HME s'ha traduït, en el context de l'anàlisi de sèries temporals, en una hipòtesi més restrictiva: la d'independència entre els successius increments relatius dels preus o rendibilitats històriques. Conceptualment, la independència implica que una variable X_t no té cap tipus d'influència sobre una altra variable Y_t. Matemàticament i tradicional[6] aquest fet s'ha estudiat mitjançant la funció de correlació entre les variables X_t i Y_t. També pot analitzar-se l'existència o no d'una relació entre els valors d'una variable (X_t) respecte els de la mateixa variable retardada k períodes (X_{t-k}), mitjançant la funció d'autocorrelació $\rho(k)$[7].

La hipòtesi d'independència és consistent amb la HME, tot i que refusar la primera no implica fer el mateix amb la segona. Segons la HME el preu d'un actiu, donat que reflexa el seu valor intrínsec, només pot variar com a conseqüència de l'arribada de nova informació al mercat. Però un cop s'ha generalitzat i els agents econòmics han reaccionat per tal d'ajustar novament el preu a un valor diferent, la informació ja no tindrà més influència en la formació futura dels preus de l'actiu. Traduït en termes d'independència, la funció d'autocorrelació dels preus d'un actiu financer tendeix ràpidament a zero a mesura que augmenta el nombre de retards considerats.

L'acceptació de la HME feu que en una primera aproximació al comportament dels mercats financers s'utilitzessin models que compleixen amb la hipòtesi d'independència. A més a més, sol acceptar-se que els canvis successius en el preu o cotització d'un actiu segueixen idèntica distribució de probabilitats. Conjugant ambdós

5 Per posar un exemple de les relacions entre els factors que influeixen en la "racionalitat" dels individus i la memòria-dependència podem mencionar el comportament dels agents davant el risc. Segons Peters és una de les causes que expliquen que el comportament del mercat en conjunt no s'acosti a l'equilibri, mentre que Poterba i Summers (1988:53) associen la dependència en els moviments d'un actiu amb el risc del mateix: "***Si els moviments en els preus dels actius contenen grans components transitoris, llavors per a inversors amb horitzons temporals a llarg termini el mercat seria menys arriscat del que semblaria quan la variància de la rendibilitat d'un període és extrapolada utilitzant el model de passeig aleatori***". [Traduït al català].

6 Si bé tradicionalment s'ha emprat la funció de correlació, actualment es disposa d'altres instruments per a mesurar la dependència tan lineal com no lineal entre els valors d'una sèrie temporal. Entre les més comuns es troben: la τ de Kendall, la ρ de Spearman, la γ de Gini, la β de Blomqvist, la σ de Schweizer i Wolf o l'índex de dependència $\phi^2_{X,Y}$. Vegi's Nelsen, Roger B. (1999:125-182).

7 A aquesta funció ens remetrem més endavant per donar una definició de procés amb memòria a llarg termini.

factors s'obté que el model òptim per a l'anàlisi de sèries temporals financeres idènticament i independent distribuïdes (IID) és el de passeig aleatori (*random walk*). En paraules d'E. Fama (1965:90):

> "E*n base a tots aquests tests es conclou que la suposició d'independència del model de passeig aleatori sembla ser una adequada descripció de la realitat.*" [Traduït al català]

Formalment, el model de passeig aleatori es representa:

$$f\,(x_{t+1}|\,\Phi_t) = f\,(x_{t+1})$$

on x_{t+1} és el valor de la variable aleatòria en el moment t+1, f és la funció de densitat de la distribució i amb Φ_t es representa un concepte tan vague com el conjunt d'informació disponible en t i de rellevància per a la variable X.

Com es pot deduir de la seva expressió formal, adoptar el model de passeig aleatori en la sèrie temporal de cotitzacions d'un actiu financer implica que el canvi en el preu futur de l'actiu és independent dels canvis que s'han donat en el passat, i en conseqüència no és possible trobar patrons històrics a partir de dades anteriors que millorin els resultats que s'obtindrien com a mitjana en el conjunt del mercat. Aquest fet és contrari a la concepció de l'anàlisi tècnica que, recolzada en la hipòtesi de memòria dels mercats financers, té com a objectiu la detecció de figures repetitives en el gràfic de sèries de preus històrics per tal de dissenyar una estratègia amb els diferents actius que superi la de "comprar i mantenir". O dit d'una altra manera, si es compleix la hipòtesi d'independència per a la variable cotització d'un actiu financer, l'estratègia de considerar que el preu el dia següent coincidirà amb el d'avui no es pot batre amb cap altra, en el sentit que l'error comès en la predicció usant qualsevol altre mètode serà sempre, a llarg termini, superior. Així, només que es compleixi la versió dèbil de la HME, l'anàlisi tècnica perd teòricament tota la seva utilitat. Si s'accepta la HME no en la forma dèbil sinó en la versió semi forta, conclouríem que tampoc l'anàlisi fonamental pot ser utilitzada per a millorar el nostre coneixement sobre el comportament futur dels mercats financers, ja que tota la informació pública ha estat recollida ja pel mercat.

Cal puntualitzar, però, que el model de passeig aleatori no diu que la informació passada no sigui significativa per a valorar una distribució futura del preu d'un actiu financer. De fet, el millor estimador, tal com s'ha comentat, per al valor futur de la variable en t+1 a manca de nova informació és el propi valor en t, sempre sota el supòsit d'estacionarietat en la sèrie temporal que recull l'evolució de la variable. El que afirma aquest model és que els canvis futurs no depenen dels canvis anteriors d'una manera seqüencial o cronològica, i que per tant no és possible trobar patrons susceptibles de repetir-se. Seguint el raonament de Fama, donat que no hi ha arguments per a defensar que la informació arriba al mercat de forma sistemàtica i no aleatòria, tampoc hi ha arguments per a defensar una naturalesa no aleatòria dels canvis en els preus.

Insistim en que el model de passeig aleatori és només una possibilitat per a recollir l'evolució del preu d'un actiu que formi part d'un mercat que segueixi la HME, però no és l'única. Abandonar per motius empírics el *random walk* no és motiu suficient per a refusar l'eficiència del mercat. Tot i que sí col·lapsarien moltes anàlisis quantitatives que es duen a terme en el context dels mercats financers, com ara el CAPM o el concepte de risc com a desviació estàndard o volatilitat[8].

3.2.3. Trencament amb la hipòtesi d'independència.

L'accepció de memòria-dependència és la que sorgeix com a crítica a la hipòtesi d'independència entre els valors de l'increment en el preu relatiu d'un actiu al llarg del temps.

En un primer pas, de l'observació del comportament dels mercats financers s'extreu que la reacció dels agents econòmics a la nova informació no és instantània. Això implica que durant el període de temps en que es produeix l'ajust entre preu i valor intrínsec, els canvis successius de la variable estan relacionats. En un entorn com l'actual, on Internet i les noves tecnologies permeten un contacte gairebé ininterromput amb qualsevol font d'informació a nivell nacional i internacional, i on la negociació electrònica possibilita les operacions en el mercat de forma gairebé contínua i molt àgil, podria semblar difícil a priori que aquest biaix temporal es produís.

Les raons que expliquen l'existència de l'esmentat lapsus de temps són principalment l'espera d'una tendència i un cert comportament de col·lectiu. La primera raó queda justificada per l'existència de costos de transacció, com ara comissions de compra i venda, que eviten ajustos automàtics per petites oscil·lacions en el preu d'un actiu

[8] Si s'utilitzen altres models diferents al de passeig aleatori per a modelitzar una sèrie temporal d'una variable financera, i es considera la possibilitat de dependència entre observacions, podem obtenir valors de la variància que invalidin la interpretació d'aquest paràmetre com a mesura del risc de l'actiu o del mercat. Així, ens podríem trobar amb variàncies infinites o bé en el cas que es perdés la linealitat: "*Una propietat important del passeig aleatori X_t és que la variància dels increments és lineal en l'interval d'observació. És a dir, la variància de X_t-X_{t-2} és dues vegades la de X_t-X_{t-1}*". Lo, A.W. i A. C. MacKinlay (1988:44) [Traduït al català].

financer, factor al que podem afegir l'interès dels inversors per aprofitar les avantatges fiscals de la realització de pèrdues o manteniment de posicions. La reacció es produirà quan la nova informació sigui indicadora d'una tendència clara. També pot ser degut a la diferent naturalesa dels inversors. Mentre aquells agents més agressius, financerament parlant, o amb horitzons d'inversió a curt termini reaccionaran ràpidament a qualsevol informació nova que es produeixi, els agents que tinguin períodes d'inversió més llargs o que per norma adoptin una actitud conservadora en front a canvis, prendran les decisions per actuar un cop hagi transcorregut més temps des de que es donà a conèixer la notícia. El segon factor fa referència a la sobrereacció en l'evolució d'una cotització com a conseqüència del comportament seguidor[9], força usual dins la pràctica financera actual. El fet de que no tothom disposi de la mateixa qualitat d'informació ni del mateix coneixement per a tractar-la porta a un sector del mercat a imitar els moviments d'agents més experimentats. Això provoca que determinades decisions influeixin en preses de decisions alienes. Tot i amb això, les correlacions degudes a aquests factors duren un període de temps breu i només explicarien la presència de dependència a curt termini.

Amb l'anterior exposició hem volgut puntualitzar que la hipòtesi d'independència, que es continua mantenint en molts models dels que s'utilitzen per analitzar el comportament d'un actiu financer, si bé es contrasta (com veurem posteriorment en el capítol) mitjançant càlculs matemàtics i estadístics, els resultats als que s'arribarien estarien buits de contingut si no es pogués donar una interpretació econòmic-financera dels mateixos.

Finalment, hi ha dos aspectes importants als que no hem fet referència però que no hem oblidat: la diferència qualitativa entre la influència d'un valor passat a curt termini o a llarg termini, de la que en tractarem en el següent apartat, i la consideració d'altres factors no estrictament quantitatius (culturals, psicològics,...) com a factors susceptibles d'influir en la valoració futura d'una variable. Aquests darrers seran introduïts en el capítol cinquè.

[9] A aquest comportament, entre altres causes, es refereix George Soros (1999:83) per denunciar la inestabilitat dels mercats financers.

3.3. Definicions de memòria en les sèries temporals.

No hi ha una definició única. Contemplada com a característica d'una sèrie temporal, és possible emprar diversos criteris per tal de distingir quina és la condició que ha de complir una sèrie per dir si presenta o no memòria, en el sentit de memòria-dependència, i en cas afirmatiu, per diferenciar entre el curt i el llarg termini. Es poden trobar classificacions que prenen com a criteris propietats basades en la funció d'autocorrelació, en la densitat espectral, en els errors de predicció, etc.

En primer lloc donarem una visió molt general i després entrarem al detall d'algunes definicions més formals, entenent que en funció de l'objectiu per al que es requereixi l'anàlisi de la dependència pot ser més adequada una definició o una altra.

3.3.1. Una definició intuïtiva.

Quan ens demanem per la memòria d'una sèrie temporal, ens estem demanant si el fet que una variable prengui un valor concret en el dia d'avui, o que el prengués en el passat, tindrà algun efecte sobre el valor que assolirà en el futur. És a dir, si hi ha o no relació o influència entre realitzacions separades en el temps. La influència, matemàticament, pot presentar formes específiques diverses, com ara l'autocorrelació, concordança, dependència quadrant positiva o negativa, a las que s'anomena de forma genèrica amb el terme 'dependència'. La qüestió que cal plantejar-se després és fins quan perduraran aquests efectes, si és que existeixen.

En la tesi doctoral d'Ernest Pons (1998:4) trobem una aportació que per a nosaltres resulta molt oportuna:

> "*A nivell intuïtiu es pot interpretar aquest concepte* [la memòria] *com la màxima distància temporal a través de la qual es manté una certa dependència entre els valors de la variable.*"

De la definició[10] anterior destaquem dos aspectes. En primer lloc, la dificultat que l'anàlisi economètrica ha trobat per a traduir en termes matemàtics el concepte de

[10] Encara que a nivell intuïtiu, la definició ens aproxima al concepte de memòria-dependència, però mostra ja les grans diferències amb les altres accepcions. No es pot fer menció a l'interval de temps màxim en el que es repeteixen les figures en la memòria-patró doncs es suposa que aquestes són estables, i tampoc podem extrapolar aquesta idea a la memòria col·lectiva, on més que la distància temporal entre

dependència entre observacions. D'altra banda, la importància donada des d'aquesta perspectiva a la duració de la memòria per sobre d'altres factors com ara la seva naturalesa, les variables que han intervingut en el manteniment dels records, etc. És per aquests dos motius que trobem una varietat de classificacions de la memòria, sempre distingint entre memòria a curt i a llarg termini. Algunes d'elles les resumim a continuació.

3.3.2. Definicions i classificacions de memòria-dependència.

De forma general i malgrat el ventall de criteris per establir la pertinença d'una sèrie temporal a un grup o altre, partirem de la generalitzada classificació de memòria que també es troba a E. Pons (1998:34) i que distingeix entre dos tipus principals de dependència en una seqüència d'observacions: (a) *memòria permanent* i (b) *memòria transitòria*. En el primer cas, una alteració en un valor concret de la variable produeix efectes en totes les observacions futures, mentre que en el segon cas, els efectes s'atenuen si bé poden perdurar més o menys en el temps. Sense ser-ne una total equivalència, s'associa aquesta distinció amb la que separa les sèries temporals no estacionàries de les estacionàries, degut precisament a que en les darreres hi ha certes característiques que es mantenen invariables a pesar d'una translació temporal.

Formalment, direm que un procés estocàstic $X = \{x_t$, amb $t \in \tau\}$ és *estacionari* si l'estructura probabilística per a les realitzacions $(x_{t1}, x_{t2}, ..., x_{tn})$ és la mateixa que per a $(x_{t1+k}, x_{t2+k}, ..., x_{tn+k})$ per a qualsevol valor enter que prengui k.

A nivell pràctic, i donada la dificultat per a que un procés compleixi amb l'estacionarietat estricta que la condició anterior comporta, es defineix una segona categoria de processos anomenats *estacionaris de segon ordre* o *estacionaris en sentit dèbil*, als quals tan sols se'ls exigeix el compliment de dues propietats:

(1) que l'esperança matemàtica de la sèrie sigui invariant respecte a translacions en el temps: $E[x_t] = \mu, \forall t$;

esdeveniments que han modificat la memòria del mercat, el que interessa és el propi procés de canvi o aprenentatge experimentat.

(2) que la funció d'autocovariances sigui invariant respecte a translacions en el temps. És a dir, la funció d'autocovariances únicament depengui del temps intern o diferència entre els instants que es consideren:
$\gamma (t, s) = E[(x_t - \mu)(x_s - \mu)] = \gamma (0, t - s) = \gamma (t - s), \ \forall t, s.$

Nosaltres ens referirem a processos estacionaris entenent els estacionaris de segon ordre.

Dins de cadascun d’aquests dos blocs principals, es troba una segona divisió. Es distingeix entre (a1) *memòria permanent infinita* i (a2) *memòria permanent finita*, per un costat, i (b1) *memòria a llarg termini*, (b2) *memòria a curt termini* i (b3) *sense memòria*, pel que fa a la subdivisió dins la memòria transitòria. Nosaltres només destacarem que els processos amb memòria permanent finita es diferencien dels de memòria permanent infinita en que els primers presenten reversió a la mitjana, i que encara que en algunes classificacions (b3) es considera un cas particular de (b2) creiem que, conceptualment, cal diferenciar-los. Al nostre entendre la diferència entre la no presència de memòria i la seva existència a curt termini és prou important com per a merèixer dos subgrups diferenciats.

Així doncs, la classificació anterior es resumeix en la taula 3.1.:

Sèries temporals	Memòria permanent	infinita
		finita
	Memòria transitòria	a llarg termini
		a curt termini
		sense memòria

Taula 3.1. Sèries temporals segons la memòria-dependència.
Font: Pons Fanals, E. (1998:9).

*

A continuació, i partint del capítol 3 de Pons (1998), elaborem la taula 3.2. on resumim les possibles definicions i tipologies de la memòria-dependència en funció del criteri escollit o propietat que ha de complir la sèrie temporal.

Classificacions de la memòria-dependència							
	en funció de l'estructura de correlacions			en funció sumes parcials	en funció de la densitat espectral	
Criteri	$A=\frac{1}{T}\sum_{k=1}^{T}\rho^2(k)$	$B=\sum_{k=1}^{\infty}\lvert\rho(k)\rvert$	$C=\sum_{k=-\infty}^{\infty}\lvert\gamma(k)\rvert$	$D=\frac{1}{T}S_T^2$	$\upsilon=\max_\omega f(\omega)$ - $\min_\omega f(\omega)$	δ, *índex del grau de memòria*
Memòria a llarg termini	$\lim_{T\to\infty} A \neq 0$	B=∞	C és convergent	$\lim_{T\to\infty} E(D)=0$ ó $\nexists$ lim	υ infinit	$\delta \neq 0$
Memòria a curt termini	$\lim_{T\to\infty} A = 0$	B<∞	C és divergent	$\lim_{T\to\infty} E(D) \neq 0$	υ finit	$\delta = 0$
Sense memòria	A=0	B=0	C=0		υ = 0	
Autors	Wiener	McLeod i Hipel	Hosking i Ding	Rosenblatt	Parzen	Parzen
Observacions	ρ(k) és f. d'autoco-rrelació per k períodes	ρ(k) és f. d'autoco-rrelació per k períodes	γ(k) és f. autocovari-ància per a k períodes	$S_T=\sum_{t=1}^{T}x_t$	f(ω) és la densitat espectral de la sèrie	$f(\omega)=\omega^{-2\delta}L(\omega)$

Taula 3.2. Classificacions de la memòria-dependència.

Es pot treballar paral·lelament en el domini temporal o en el domini de les freqüències. El nexe entre un i altre és observable mitjançant la relació que es pot establir entre la funció d'autocovariància i la densitat espectral a través de la igualtat:

$$f(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k e^{-ik\omega}=\frac{1}{2\pi}\gamma_0+\frac{1}{\pi}\sum_{k=1}^{\infty}\gamma_k\cos(k\omega)$$

Si en el domini temporal una sèrie presenta memòria a llarg termini quan la funció d'autocorrelació decreix hiperbòlicament ($\rho(k)=(1+k)^{\delta-1}$;, $0\leq\delta\leq1$), mentre que serà a curt termini quan el decreixement és exponencial ($\rho(k)=\gamma^k$; $0\leq\gamma\leq1$); des del punt de vista del domini de les freqüències, un procés presenta memòria a llarg termini si la densitat espectral creix sense límit quan la freqüència tendeix a zero.

*

Amb tot aquest ventall de tipologies[11] hem volgut remarcar únicament l'interès que s'ha mostrat des de l'econometria per l'estudi de conjunts de dades on no es compleix la hipòtesi d'independència entre observacions, sent necessària la recerca de definicions i mesures per a determinar el grau d'influència de valors passats en el comportament futur de la variable.

Els criteris presentats a la taula 3.2. no exhaureixen la possibilitat de classificar una sèrie temporal amb memòria a llarg o a curt termini mitjançant l'estructura de correlacions. A. W. Lo (1991:1286), amb importants treballs empírics sobre l'anàlisi de la dependència en variables financeres, empra una definició alternativa. C.W.J. Granger i Z. Ding (1996:261) descriuen una sèrie temporal amb *memòria a llarg termini* com aquella que té un correlograma (la gràfica de la funció d'autocorrelació estimada entre la variable X_t i la variable X_{t-k} versus el valor de k) que descendeix lentament cap a una asímptota horitzontal a y=0.

Tot i que no hi ha unanimitat en el criteri a utilitzar per a classificar una sèrie temporal en funció de la dependència entre observacions, l'estructura de correlacions i la densitat espectral són els més emprats. Dins del primer tipus és possible basar-se en els valors de la funció d'autocovariància o de la d'autocorrelació. Donat que per a arribar a $\rho(k)$ únicament cal dividir $\gamma(k)$ per la variància de les dades que configuren la mostra, la conclusió a la que s'arribi no es veurà afectada per l'ús d'una o altra, sempre i quan $\gamma(0)$ sigui finita. La preferència per la funció d'autocorrelació que mostren alguns autors s'explica per la voluntat d'homogeneïtzar resultats de variables amb diferent dispersió.

Acceptant que les sèries temporals de variables econòmiques i financeres poden caure dins el grup de variables amb un comportament que assenyala presència de memòria[12], el següent pas a donar és cercar el model més adient per a recollir i explicar la seva evolució.

A continuació, farem un repàs dels principals models proposats des de l'àmbit de l'econometria. L'existència d'altres treballs que tracten amb suficient amplitud i profunditat aquests models i el nostre interès en dur l'anàlisi cap a les XNA, ens permet limitar la descripció als que més comunament s'han utilitzat per a l'estudi de sèries temporals de preus i rendibilitats de diferents mercats financers i que han entrat en el debat sobre la naturalesa de la memòria en ells.

[11] Encara es poden apuntar altres classificacions, derivades de l'aplicació de l'anàlisi de sèries temporals al càlcul de prediccions i del concepte d'informació, entre memòria a llarg i a curt termini que no detallem aquí [vegi's Pons, E. (1998:83 i ss.)].

[12] Sol acceptar-se la independència entre observacions d'una variable financera per motius de necessitat operativa, però després de la recerca realitzada sobre aquest tema és usual considerar l'existència de dependència en relació a valors històrics. Així, E. Panas (2001:395) comenta: "*Està àmpliament acceptat que les rendibilitats dels actius (1) segueixen una distribució de cues pesades no normal, (2) tenen autocorrelacions i autocorrelacions parcials que no decauen ràpidament a zero, i (3) semblen tenir cicles no periòdics*".

3.4. Modelització de la memòria a llarg termini.

3.4.1. Models per a sèries temporals amb memòria.

L'existència de memòria en una sèrie temporal ha impulsat la recerca de models més complexos que el de passeig aleatori, models adients per a explicar el comportament d'aquelles variables on les observacions es veuen influïdes pels valors passats. La forma concreta que prendrà cadascuna de les modelitzacions estarà en funció de la memòria que s'hagi detectat, tot i que per a cada categoria són diversos els models que poden emprar-se, tenint tots avantatges i inconvenients que hauran de tenir-se en compte a l'hora de fer l'elecció.

De tots els casos en els què ens podem trobar, la modelització de sèries temporals amb memòria a llarg termini ha gaudit d'especial atenció per part dels investigadors. Dues raons ho justifiquen. En primer lloc, la important quantitat de sèries temporals de variables que en presenten, entre d'altres i d'interès particular per a nosaltres, les econòmiques i financeres. I per altre costat, si bé els processos ARMA i ARIMA modelitzen de forma adequada les sèries amb memòria a curt termini i permanent, respectivament, les sèries amb memòria a llarg requereixen models[13] intermedis més complexos, amb dificultats pràctiques per a l'estimació de paràmetres i que comporten la recerca d'alternatives més viables.

Hi ha una important relació entre la predicció com aplicació de l'anàlisi de sèries temporals i la memòria-dependència. El nostre objectiu no és analitzar aquesta última per millorar la predicció de cotitzacions futures en un mercat financer, però si ho fos, interessaria trobar quanta més dependència entre observacions millor, ja que quan més gran sigui tant el grau de dependència com el nombre de retards durant els quals es manté la relació, més gran serà l'interval temporal durant el que es podran realitzar prediccions acceptables. La nostra aportació és metodològica i consisteix en, d'una banda, proposar una nova forma d'apropar-nos a l'obtenció i interpretació de la dependència entre els valors d'una sèrie temporal a través de l'anàlisi de les matrius de pesos d'una XNA, i d'altra, comparar aquesta influència amb els efectes del passat deguts a la memòria col·lectiva, emprant també en aquest cas els resultats obtinguts

[13] Malgrat existeixen i ja s'han emprat alternatives a la modelització merament lineal de la dependència, per motius de simplicitat són aquests els que més s'utilitzen en l'anàlisi de dades financeres i és a ells únicament als que ens remetrem en la tesi. Exemples de models capaços de recollir la dependència no

en la matriu de pesos un cop entrenada la xarxa amb dades històriques. A aquest aspecte hi tornarem en l'apartat 3.7.

3.4.2. L'anàlisi R/S.

D'entre els diferents models per a detectar i analitzar la memòria a llarg termini d'una sèrie temporal, començarem destacant el pioner anàlisi R/S desenvolupat per H.E. Hurst.

L'hidròleg britànic Harold Edwin Hurst viatjà a Egipte per col·laborar en la construcció d'una presa al riu Nil. L'objectiu del seu viatge era determinar el nivell òptim de reserves que havia de mantenir l'esmentada presa. Però els resultats que obtingué en els seus estudis feren que el treball d'unes quantes setmanes passés a ser una investigació a la que Hurst hi dedicaria en torn a uns 40 anys. La "culpable" del canvi és la que es coneix com *anàlisi dels rangs reescalats* o *anàlisi R/S* (en terminologia anglesa, *rescaled range* o bé *range over standard deviation* fent menció al seu procediment de càlcul).

El nivell òptim per a la presa havia de ser suficient per tal d'evitar que una acumulació excessiva d'aigua en un període pogués produir els usuals desbordaments del Nil, però tenint en compte alhora que havia de cobrir una quantitat mínima destinada a garantir la disponibilitat d'aigua en tot moment per al conreu i per al consum de la població (uns elevats costos de construcció expliquen la importància de l'estudi que se li assignà). L'aparent aleatorietat de les entrades degudes a la pluja dificultava el càlcul del volum de reserva desitjable. Hurst, però, tenia l'avantatge de disposar d'un gran volum d'informació del nivell del riu en el passat, en concret, podia comptar amb valors d'una sèrie que comprenia més de 800 anys.

De l'observació i anàlisi de les dades, Hurst s'adonà que l'ordre cronològic d'aquestes és fonamental per a trobar la capacitat adequada de la presa. Una variable en la que l'estructura temporal de les dades sigui un tret important per a l'evolució de les mateixes és indicatiu de que hi ha dependència entre les observacions[14]. En la sèrie que estudià Hurst és fàcil comprovar la rellevància de l'ordre en la successió de les

---

lineal són els ARCH, iniciats amb els treballs de Robert F. Engle, els models STAR, o la incipient teoria de còpules.

[14] Quan una variable aleatòria es caracteritza per la independència entre observacions, una alteració en l'ordre de les dades continua mantenint el model aleatori inicial. Mentre que si en una sèrie temporal en la

dades[15]. Aquest fet l'induí a investigar sobre les característiques de la memòria de la sèrie, demanant-se fins a quin punt els valors futurs depenen dels anteriors i si és possible detectar la periodicitat dels cicles en la variable estudiada. L'anàlisi R/S busca un valor que mesuri el grau i sentit de la influència entre observacions d'una sèrie temporal. En definitiva, H.E. Hurst (1951) proposà un mètode per a detectar la memòria-dependència a llarg termini.

L'*estadístic R/S* es defineix com el rang de les sumes parcials de les desviacions de la sèrie temporal respecte la seva mitjana, i reescalat per la desviació estàndard. L'anàlisi fou aplicada en primer lloc a fenòmens naturals, com ara al propi nivell del riu Nil o a les taques solars. Més tard però, autors d'entre els que destaquem Benoit B. Mandelbrot (1972), i posteriorment Andrew W. Lo (1991) i Edgar E. Peters (1991 i 1994) la recuperarien i aplicarien a l'estudi de sèries econòmiques[16].

El punt de partida comú a tots els treballs on s'aplica l'anàlisi R/S és la concepció de que el model de passeig aleatori consistent amb la HME no és una bona modelització del comportament de les rendibilitats dels actius en els mercats financers. Per veure la relació entre l'anàlisi R/S i la dependència entre valors d'una sèrie temporal, detallem a continuació els passos que calen donar per trobar el valor de l'exponent de Hurst, tot seguint les indicacions d'E.E Peters (1991 i 1994).

De la sèrie temporal de valors observats $\{Z_t\}$ amb t = 1, 2, ...m, obtenim una nova sèrie $\{X_t\}$ de longitud m-1 aplicant la transformació:

$$X_t = \ln \frac{Z_{t+1}}{Z_t} \quad \text{amb } t = 1, 2, ..., m-1$$

que es dóna una relació de dependència modifiquem l'ordre cronològic de les observacions, es perd aquesta propietat i la sèrie pot convertir-se en un exemple de model de passeig aleatori.

[15] Pot fer-se amb un senzill exemple numèric. Considerem: 1- El volum anual d'aigua necessari per cobrir les necessitats de la població que abastarà la presa és constant i es fixa en 50 unitats de capacitat (u.c.). 2a- El volum d'aigua degut a la pluja i a altres fenòmens durant 8 anys ha evolucionat de la següent manera: 75, 25, 75, 25, 75, 25, 75, 25 u.c. En conseqüència, obtindríem que el nivell anual de reserva de la presa ha estat de 25, 0, 25, 0, 25, 0, 25, 0 u.c. Una presa amb una capacitat de 25 u.c. és suficient per cobrir les necessitats de la població. 2b- Considerem ara els mateixos valors per a les entrades d'aigua a la presa però en un ordre diferent: 75, 75, 75, 75, 25, 25, 25, 25 u.c. L'evolució del nivell de la presa en aquest segon cas seria ben diferent: 25, 50, 75, 100, 75, 50, 25, 0 u.c. On ara és necessari incrementar el nivell mínim de reserva de la presa a 100 u.c. per evitar un desbordament del Nil i assegurar l'abastament a la població en els anys de sequera.

[16] Yao i Tan (2000:83) defineixen l'anàlisi R/S com: "*És un estadístic robust per a mesurar la quantitat de soroll en el sistema. Pot ser emprat per a determinar la persistència i la longitud mitjana dels cicles no periòdics.*" [Traduït al català]. Els autors l'utilitzen per a detectar memòria a llarg termini en sèries temporals de tipus de canvi.

Si Z_t és el preu o cotització d'un actiu financer, la variable X_t representa el rendiment de l'actiu, en termes logarítmics[17].

La nova sèrie X_t es subdivideix en un conjunt d'A subperíodes de longitud N, amb interseccions buides (la darrera dada d'un subgrup s'exclou del període següent).

Per a cada subperíode es calcula el rang mitjançant l'expressió:

$$R_{i,N} = \max_{1 \le k \le N} (x_{i,N,k}) - \min_{1 \le k \le N} (x_{i,N,k}) \quad \text{amb } i = 1, 2, \ldots A$$

on $x_{i,N,k} = \sum_{j=1}^{k} (x_{i,j} - \mu_{i,N})$ és una mesura de la distància acumulada entre cadascuna de les observacions i la mitjana del subperíode i $\mu_{i,N}$ és la mitjana aritmètica dels valors de la variable X_t compresos en l'i-éssim subperíode de longitud N.

Amb la finalitat de comparar fenòmens de diferent naturalesa, H.E. Hurst dividí el rang anterior per la desviació estàndard de les observacions,

$$S_{i,N} = \left[\frac{1}{N} \sum_{j=1}^{N} \left(x_{i,j} - \mu_{i,N} \right)^2 \right]^{\frac{1}{2}},$$

dins de cadascun dels intervals de duració N. Això permet tractar una sèrie temporal de preus d'actius des de 1920 fins a 1980, prenem per cas, eludint els problemes derivats de la inflació.

D'aquesta manera queda definit l'estadístic $R_{i,N}$ / $S_{i,N}$ per a tot N i per a cada subperíode des d'i = 1, ..., A. El numerador és sempre positiu ja que es tracta d'un rang. El denominador és una desviació estàndard i per tant també és positiu. En conseqüència, el resultat del quocient serà sempre més gran o igual a zero.

La mitjana $(R/S)_N = \frac{1}{A} \sum_{i=1}^{A} \left(\frac{R_{i,N}}{S_{i,N}} \right)$ és el resultat de l'estadístic R/S per a cadascun dels valors de N, definits en el moment de formar els subperíodes. Així s'obté un conjunt de parells [$(R/S)_N$, N], en els que s'observa que la primera component augmenta a mesura que també ho fa l'amplitud dels intervals.

[17] Des del punt de vista d'una operació d'inversió a un període, Z_t representaria l'*input* de l'operació, Z_{t+1} l'*output* que s'obtindria, i el valor X_t resultant mesuraria el *rendiment estricte efectiu brut*, establint a la seva vegada la relació entre el *termini financer mitjà* de l'operació i la *taxa financera de rendibilitat*. [Vegi's Rodríguez, A. (1997:20 i ss.)]

La relació entre l'estadístic $(R/S)_N$ i la longitud del subperíode ve determinada per l'expressió:

$$(R/S)_N = \alpha \cdot N^H$$

on α és una constant i H rep el nom d'*exponent de Hurst*[18]. És precisament el valor d'aquest exponent el que determina, en l'anàlisi R/S, la naturalesa de la sèrie temporal estudiada. Per a trobar-lo, donat que tenim parells de nombres que relacionen l'estadístic amb la longitud dels intervals formats, podem prendre logaritmes i representar els punts en un eix de coordenades. La pendent de la recta resultant és una estimació de l'exponent de Hurst.

$$\log (R/S)_N = \log \alpha + H \log N$$

L'hidròleg establí un lligam entre la funció d'autocorrelació a llarg termini, C, i l'exponent H de la següent manera[19] :

$$C = 2^{(2H-1)} - 1$$

Tornant a la hipòtesi d'independència ja comentada, si d'una sèrie temporal podem dir que no té memòria esperaríem que fos C=0. Per tal que això succeeixi, H ha de ser proper a 0'5. Una altra interpretació possible apunta a que si la sèrie temporal segueix un model de passeig aleatori, el rang de les desviacions acumulades hauria d'augmentar en proporció a l'arrel quadrada del temps (N).

$H \neq 0'5$ indica que les observacions de la variable no són independents entre sí. Cadascuna d'elles porta incorporada una part de memòria de tots els esdeveniments que l'han precedida. És per tant, una memòria a llarg termini, que com afirma E.E. Peters (1991:64), teòricament s'hauria de conservar sempre. Tot i que esdeveniments més recents tinguessin una repercussió major que d'altres més allunyats en el temps, aquests encara haurien de conservar una influència residual. El que ha passat avui influeix en el futur. On som és el resultat d'on hem estat, d'una llarga cadena d'esdeveniments interconnectats. Aquesta mateixa idea la recollim en l'apartat 3.7. per a expressar la memòria-dependència en forma d'una combinació de valors del passat a través dels pesos d'una XNA.

[18] Hurst anomenà k a aquest exponent però, anys més tard, Benoit Mandelbrot el tornà a batejar amb la inicial del seu descobridor.

[19] Hurst coneixia els treballs que duia a terme Albert Einstein sobre el moviment erràtic d'una partícula sobre la superfície d'un fluid (moviment brownià), i aprofità els resultats per a relacionar-los amb els

Tornant a la interpretació dels valors de l'exponent de Hurst, cal però fer una important distinció entre els que són superiors i els inferiors a 0'5 dins l'interval obert (0, 1)[20].

- Si H $\in$ (0, 0'5), la funció d'autocorrelació és negativa i la sèrie temporal és *antipersistent*, també anomenada *ergòdica* o amb *reversió a la mitjana* ($\Delta x_t \Rightarrow \nabla x_{t+1}$).
- Si H $\in$ (0'5, 1), la funció de d'autocorrelació és positiva i per tant variacions en el valor de la variable van seguits de variacions amb el mateix signe[21]. De la sèrie es diu que és *persistent* o amb *tendència reforçada* ($\Delta x_t \Rightarrow \Delta x_{t+1}$). Altres noms que també fan referència a sèries temporals amb aquesta propietat són els de *moviment brownià fraccionari* (*fractional brownian motion*) o *passeig aleatori esbiaixat* (*biased random walk*).

Als nostres efectes, són d'especial interès els resultats que s'han obtingut en aquells casos en que s'ha aplicat l'anàlisi R/S en els mercats financers. El primer fet destacat és que els resultats han estat força variats, tot i que en els casos en que s'ha detectat dependència, H ha estat dins l'interval (0'5, 1)[22]. La interpretació financera que podem donar és que la informació que es rebi avui en el mercat continuarà sent descomptada en el futur, encara que només durant un interval temporal acotat.

Sense haver de recórrer a cap mena de càlcul numèric, i tal com va fer Hurst a l'inici de la seva recerca, l'observació dels mercats financers ja ens dóna indicis per a sospitar que hi ha sèries històriques de preus que no són completament aleatòries. La primera pista ens la proporciona l'existència de períodes baixistes i alcistes (els anomenats *bull* i *bear markets*) que ens mostren que les pujades i baixades en els preus dels actius financers són persistents i no només el resultat casual que fa que una variable aleatòria es mogui en una sola direcció durant un període de temps concret.

La memòria a llarg termini, malgrat que teòricament hagi de durar per sempre, és una memòria finita. És d'esperar que per a valors d'N suficientment grans l'autocorrelació tendeixi a zero, la influència del passat sigui gairebé nul·la i la sèrie convergeixi a

obtinguts amb l'anàlisi de rangs reescalats. Einstein va trobar que la distància que una partícula aleatòria cobreix augmenta amb l'arrel quadrada del temps utilitzat per a mesurar-la. $R = T^{0'50}$.

[20] El procés X_t és estacionari i invertible únicament per a valors de l'exponent de Hurst dins de l'interval obert (0, 1). D'aquí que es limiti la seva interpretació dins aquests límits.

[21] La dependència en les sèries temporals persistents és independent de l'escala temporal. Els canvis diaris estan correlacionats amb els canvis diaris passats i els canvis setmanals estan correlacionats amb els canvis setmanals passats.

H=0'5. Quan això succeeix es pot contemplar un canvi en la pendent de la gràfica que relaciona el logaritme de $(R/S)_N$ amb el logaritme de N, adoptant la pendent d'un angle de 45º. Aquest fet és signe de l'acabament d'un cicle. La longitud dels cicles de memòria varien d'un mercat a un altre igual que poden variar d'un actiu a un altre en un mateix mercat[23].

3.4.3. L'anàlisi R/S modificada.

El principal problema que presenta l'anàlisi de rangs reescalats desenvolupat per Hurst és que els resultats obtinguts no distingeixen entre els efectes de la memòria a llarg i a curt termini. Per aquest motiu, és possible que sèries temporals amb un H clarament superior a 0'5 i que es classificarien com a sèries persistents, únicament presentin dependència entre observacions durant un període temporal breu.

Una primera possibilitat per a que l'anàlisi R/S sigui un mètode vàlid per a detectar la memòria a llarg termini és eliminar abans de la seva aplicació els efectes de la memòria a curt. És a dir, com a primera aproximació pot aplicar-se el procediment descrit a l'apartat anterior però no sobre els valors de la sèrie original, sinó sobre els residus de la sèrie un cop se li han extret els efectes de la dependència a curt termini, per exemple, mitjançant un model autorregressiu d'ordre 1, un AR(1)[24].

Aquesta solució té l'inconvenient d'haver de definir prèviament una relació funcional concreta per a la dependència a curt termini, ja sigui en forma de model autorregressiu, AR(p), mitjana mòbil, MA(q), o una combinació d'ambdues, ARMA(p,q). Una segona possibilitat que evita tal concreció és la proposada per Andrew W. Lo a l'any 1991.

Prenent com a base l'estadístic R/S tradicional, Lo (1991:1289) modifica únicament el denominador i obté el que anomena *estadístic R/S modificat* (*the Modified R/S Statistic),* Q_N. *S*eguint amb la mateixa notació de l'apartat anterior, tenim

[22] Només com a anècdota, comentar que el valor que va obtenir Hurst per al nivell de l'aigua en el riu Nil fou de 0'91.
[23] No ens hem referit en detall a la formació de cicles en les variables financeres, però sí que volem destacar en aquest sentit el treball de Mandelbrot (1972) en el que utilitza l'anàlisi R/S per a la detecció de cicles no periòdics en variables econòmiques, així com el d'E.E. Peters (1992) on estudia els cicles en actius de diferent naturalesa al mercat nord-americà.
[24] Si la variable X_t es distribueix segons un procés autoregressiu d'ordre 1, llavors $X_t = \rho X_{t-1} + \varepsilon_t$, on $|\rho| \in (0, 1)$ i $\varepsilon_t \sim WN(0, \sigma^2)$ és un soroll blanc.

$$Q_{i,N} = \frac{1}{\sigma_{i,N}(q)}\left[\max_{1\le k\le N}(x_{i,N,k}) - \min_{1\le k\le N}(x_{i,N,k})\right]$$

on $x_{i,N,k}$ es defineix de la mateixa manera que a l'estadístic R/S tradicional. Q_N és, llavors, la mitjana aritmètica dels A estadístics $Q_{i,,N}$ obtinguts. La desviació estàndard de les observacions es substitueix per:

$$\sigma_{i,N}(q) = \frac{1}{N}\sum_{j=1}^{N}\left(x_{i,j} - \mu_{i,N}\right)^2 + \frac{2}{N}\sum_{j=1}^{q}\omega_j(q)\left(\sum_{k=j+1}^{N}\left(x_{i,k} - \mu_{i,N}\right)\left(x_{i,k-j} - \mu_{i,N}\right)\right).$$

Si la funció d'autocovariància l'escrivim com $\gamma_{i,j}$ i la variància de l'interval com $\sigma^2_{i,N}$, s'obté,

$$\sigma_{i,N}(q) = \sigma^2_{i,N} + 2\sum_{j=1}^{q}\omega_j(q)\gamma_{i,j}$$

on els pesos[25] $\omega_j(q)$ venen donats per la diferència $1 - \frac{j}{q+1},\ q < N$.

La major dificultat que presenta aquesta anàlisi és la no existència de cap mètode per a determinar quin és el valor idoni per al paràmetre q. Sol emprar-se la proposta realitzada per Andrews (1991) i seguida per Lo en el seu article,

$$q = [k_n] \quad i \quad k_n = \left(\frac{3n}{2}\right)^{\frac{1}{3}}\left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{\frac{2}{3}}$$

on $[k_n]$ és el més gran dels enters inferiors o igual a k_n i on $\hat{\rho}$ és l'estimador del coeficient d'autocorrelació de primer ordre de les dades.

Si la sèrie temporal estudiada correspon a una variable amb observacions independents dels valors del passat, l'estadístic R/S modificat serà igual a 0'5. Presenta com a aventatges que pot distingir entre dependència a curt i a llarg termini mesurant únicament la darrera, mentre que l'estadístic R/S tradicional no pot fer-ho, i que és igualment aplicable a sèries no normals i heterocedàstiques[26] sense que això desvirtuï els resultats obtinguts.

[25] La definició dels pesos és la suggerida a Newey i West (1987). D'aquesta manera, i permetent que q augmenti amb el nombre d'observacions N, el denominador de $Q_{i,\,N}$ s'ajusta adequadament a les formes generals de dependència a curt termini.

[26] Els models heterocedàstics són emprats per a modelar sèries temporals amb una volatilitat no constant en el temps. Tanmateix, entre altres característiques com les de predir variàncies futures o recollir en un model teòric variàncies que puguin manifestar-se en forma d'ones, també inclouen la identificació de l'existència de memòria en el procés.

Així, Lo va trobar que en estudis sobre sèries temporals financeres en les que s'havia utilitzat el valor de H com a tret distintiu per a discernir si la sèrie tenia memòria a llarg termini o no, aplicar l'estadístic R/S modificat podia afectar de tal manera que el que es pensava que era memòria a llarg termini era tan sols una forta dependència a curt. Com a norma general, els valors de l'estadístic modificat són menors als de l'estadístic tradicional i per tant l'afirmació de dependència a llarg termini basant-se en aquest mètode no és tan asseverativa com en els casos en els que s'empra l'estadístic de Hurst.

No obstant, són molts els estudis en els que es calcula l'exponent de Hurst com a una dada de referència. En primer lloc, perquè el seu càlcul és fàcilment computable mitjançant l'algorisme que el descriu i en segon lloc, perquè forma part d'un dels estimadors del paràmetre 'd' en els molt emprats models ARFIMA.

3.4.4. Els models ARFIMA.

És, sens dubte, la modelització més coneguda i utilitzada per a les sèries temporals que presenten memòria a llarg termini. De forma molt resumida, els models ARFIMA són la generalització dels processos ARIMA, deguda a C.W.J. Granger i R. Joyeux (1980) i a J.R.M. Hosking (1981), permetent que l'ordre de diferenciació sigui un valor no enter.

Traduint literalment a Pynnönen i Knif (1998:258):

> "*Tradicionalment, l'ordre d'integració més comú analitzat en sèries temporals econòmiques és zero o u. Una sèrie amb ordre d'integració zero té una variància finita independent del temps. Aquestes sèries són estacionàries en el sentit que una pertorbació només té un efecte temporal en les futures realitzacions de la sèrie i s'espera que la sèrie creui la seva mitjana dins un període de temps finit. (...) Una sèrie amb ordre d'integració u, d'altra banda, és no estacionària en el sentit que la variància s'acosta a infinit amb el temps. Una pertorbació o innovació té un impacte permanent en futures realitzacions de la sèrie i, així, el temps esperat per al retorn a la mitjana és infinit*."

És a dir, en el món econòmic i financer s'acostuma a treballar amb sèries temporals estacionàries. Donat que les dades reals sovint no s'ajusten a aquesta hipòtesi, el que es fa és diferenciar-les tants cops com sigui necessari fins a obtenir la sèrie estacionària desitjada. Els models ARIMA(p, d, q) (autoregressius d'ordre p, i mitjana mòbils d'ordre q, amb integració d'ordre d) realitzen tal funció, sempre i quan, es diferenciïn les dades un nombre enter de vegades. Formalment, un model ARIMA(p, d, q) es pot representar com:

$$(1-\phi_1 L-\ldots-\phi_p L^p)(1-L)^d X_t=(1-\theta_1 L-\ldots-\theta_q L^q)\,\varepsilon_t$$

o bé $$\Phi(L)\nabla^d X_t=\Theta(L)\,\varepsilon_t$$

on ε_t representa una variable soroll blanc, d és zero (llavors el model és un ARMA) o un nombre natural, ∇^d és l'operador diferenciació[27] i els polinomis $\Phi(L)$ i $\Theta(L)$ tenen les arrels fora del cercle unitat.

Amb els models ARIMA podem representar el comportament de sèries temporals amb memòria permanent infinita (seguint la classificació establerta a l'inici del capítol, vegi's pàgina 76), i els models ARMA recullen l'evolució de les sèries amb memòria transitòria a curt termini. Tot i així, la dificultat es trobava en els casos intermitjos, això és, en la modelització de sèries amb memòria transitòria a llarg termini i amb memòria permanent finita. Calia una família de models que cobrís aquest buit. J.R.M. Hosking s'adonà que si als models ARIMA se'ls permetia graus de diferenciació no enters, s'obtenia una modelització per a les sèries que presentaven una lenta[28] disminució de la funció d'autocorrelació a mesura que augmentava el retard entre les seves observacions. Tal com el mateix autor exposa (1981:165), el seu objectiu era trobar una manera de modelar explícitament la dependència o persistència a llarg, que fos suficientment flexible per explicar ambdós, tant l'estructura de correlació de la sèrie a curt com a llarg termini, i que permetés una fàcil generació de sèries a partir del model.

Un cop definit el model ARIMA, el pas als ARFIMA (*AutoRegressive Fractally Integrated Moving Average*) és immediat.

Un procés X_t segueix un model ARFIMA(p, d, q) si X_t és una solució de l'equació estocàstica en diferències:

[27] L'operador diferenciació pot expressar-se en termes de l'operador retard L mitjançant la següent expansió binomial: $\nabla^d=(1-L)^d=\sum_{j=0}^{\infty}\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}L^j$. Per definició $LX_t =_{def} X_{t-1}$, i per tant ∇^d = $(I-L)^d$.

[28] Quan es parla d'una lenta disminució en la funció d'autocorrelació s'està fent referència a un decreixement de tipus hiperbòlic, en concret, se sol suposar que: "asimptòticament les autocorrelacions es poden representar segons $\left|\gamma_k\right| \iota\ k^{2\delta-1}$ amb $0\leq\delta\leq 1/2$. Recordi's que en models ARMA el decreixement de les autocorrelacions segueix un esquema exponencial que es pot representar com $\left|\gamma_k\right| \iota\ C^k$." Pons, E. (1998:25).

$$\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t$$

on $\varepsilon_t, \Phi(L)$ i $\Theta(L)$ es defineixen igual que en el cas dels ARIMA(p, d, q) amb la diferència que ara $d \in$(-0'5, 0'5), essent p i q nombres enters.

Els ARFIMA permeten modelar la dinàmica tant de curt com de llarg termini d'una sèrie temporal. Així,

> "*el comportament a curt termini està recollit pel biaix dels paràmetres autoregressius i mitjana mòbils usuals mentre que la dinàmica a llarg termini està representada per un únic paràmetre: el paràmetre d'integració fraccional d*."[29]

De forma general, 'd' pot prendre qualsevol valor. La limitació establerta dins l'interval (-0'5, 0'5) és únicament per assegurar que el procés $\{x_t\}$ obtingut a partir del model sigui tant estacionari com invertible. Per exemple, quan la sèrie presenti memòria permanent finita, el paràmetre 'd' dels models ARFIMA es trobarà fora d'aquest interval. La taula[30] 3.3. resumeix les característiques de la sèrie en funció de 'd' i de l'exponent de Hurst.

Memòria	Estacionarietat	Model	d	H
Transitòria/Curta	Sí	ARFIMA	(-0'5, 0)	(0, 0'5)
Transitòria/Curta	Sí	ARMA	0	0'5
Transitòria/Llarga	Sí	ARFIMA	(0, 0'5)	(0'5, 1)
Permanent/Finita	No	ARFIMA	0'5	1
Permanent/Finita	No	ARFIMA	(0'5, 1)	(1, 1'5)
Permanent/Infinita	No	ARIMA	1	1'5
Permanent/Infinita	No	ARFIMA	(1, ∞)	(1'5, ∞)

Taula 3.3. Paràmetres de les sèries temporals amb memòria.

A continuació interpretem breument els valors més significatius del paràmetre 'd', del qual es diu que mesura el *grau de memòria d'una sèrie temporal*, tot i que la seva traducció és automàtica gràcies a la correspondència amb l'exponent de Hurst (vegi's pàg. 84).

[29] Lardic, S i V. Mignon (1999:41). [Traduït al català]
[30] Font: Pons Fanals, E. (1998). La columna de l'exponent de Hurst és afegida per part nostra.

- Si $d \in (-0'5, 0)$, la sèrie temporal és *antipersistent*. Això vol dir que les autocorrelacions són negatives i, a més a més, disminueixen (en valor absolut) seguint una funció hiperbòlica i monòtona cap al valor zero.
- Si d=0, el procés no té memòria o té memòria a molt curt termini, és un *soroll blanc*, amb correlacions molt properes a zero.
- Si $d \in (0, 0'5)$, la sèrie temporal té memòria a llarg termini i és *persistent*, és a dir, amb valors de la funció d'autocorrelació positius i que segueixen la forma $\gamma_k \sim \lambda k^{-\alpha}$, amb $\lambda > 0$, $0 < \alpha < 1$, quan $k \to \infty$. En aquest cas, a més, es pot dir[31] que $\alpha = 1-2d$.

Un dels grans problemes amb els que es troba l'ús dels models ARFIMA és l'estimació de l'ordre d'integració fraccional 'd'. En la taula 3.3. és fàcil comprovar com una primera i directa estimació sorgeix de la relació d=H-0'5. A més hi ha tot un ampli camp de recerca que ha tractat sobre l'estimació semiparamètrica, en la qual no és necessari especificar un model concret per al procés que genera les dades d'estudi, sinó que únicament considera uns supòsits sobre les seves característiques, tot i que no n'entrarem a discussió. Únicament destacarem el mètode d'estimació proposat per Geweke i Porter-Hudak (1983) degut a la seva àmplia difusió en l'anàlisi de les variables econòmiques i financeres. El caracteritza la seva simplicitat, però té l'inconvenient que no permet estimar simultàniament el paràmetre 'd' que mesura la memòria a llarg termini, amb els paràmetres 'p' i 'q' associats a la dependència a curt termini, amb la conseqüent possibilitat de distorsions en la modelització final de les observacions.

*

Concloem aquí l'estat de la qüestió sobre el tractament de la memòria des de la vessant de l'econometria, tot i ser conscients de que l'anàlisi de sèries temporals és molt més ampli i que obviem una gran quantitat de problemes i solucions aportades des d'aquest camp. Per exemple, els models ARFIMA no es comporten adequadament quan la sèrie temporal a analitzar presenta components estacionals, on es proposa l'ús de models alternatius, com poden ser els GARMA o els SARFIMA. També s'han desenvolupat models heterocedàstics com els ARCH i GARCH que contemplen els casos en els que la variància de la sèrie no és constant, sinó que varia en el temps. La recerca en el camp econòmetric és molta i molt fructífera.

[31] Hosking, J.R.M. (1996:272).

Nosaltres hem volgut referir-nos a una de les vies, i la més important en extensió, amb les que es tracta el concepte de memòria dins els mercats financers. S'ha mostrat l'adequació de l'instrumental economètric per a recollir-la, sempre que es defineixi en algun dels sentits exposats en l'apartat 3.3., i dins el context de l'anàlisi de sèries temporals.

De forma complementària, l'argument que defensem aquí és que els efectes del passat no es mesuren únicament pel grau de dependència entre dues observacions separades per un interval de temps, sinó que hi ha altres factors (històrics, deguts a l'experiència dels participants, deguts a característiques particulars dels mercats, etc.) que fan que un mercat evolucioni de la manera com ho fa. Aquests elements, que participen de la memòria col·lectiva, els podem recollir mitjançant els pesos d'una xarxa neural artificial. Per tal de comparar els efectes d'ambdues accepcions, proposem analitzar també la memòria-dependència des de l'enfocament connexionista. Abans, comentarem els resultats de l'anàlisi de la memòria en sèries temporals emprant els models desenvolupats per l'econometria.

3.5. Memòria-dependència: Treballs empírics a favor i en contra.

Sota la perspectiva de l'anàlisi economètrica no es pot establir l'existència o no de memòria en les sèries temporals de variables financeres d'una forma general. És per aquest motiu que els treballs de recerca encaminats a la detecció de dependència entre les observacions de preus d'un actiu financer o altres variables econòmiques únicament poden extreure conclusions particularitzant per a les dades en estudi i per als instruments d'anàlisi emprats. Sota aquestes restriccions s'obtenen resultats molt variats.

Hi ha un ampli ventall[32] de documents publicats on s'han aplicat els mètodes usuals de detecció de dependència entre observacions, l'anàlisi R/S i els models ARFIMA, a dades de diferents mercats financers. No pretenem fer un llistat exhaustiu dels seus resultats, sinó únicament assenyalar la diversitat de conclusions a les que han arribat, com a mostra de la dificultat de parlar de la naturalesa de la memòria, ni en sentit estrictament de memòria-dependència, en el context financer.

[32] Berg i Lyhagen (1998), Blasco i Santamaría (1996), Cheung (1995), Cheung i Lai (2001), Henry (2002), MacDonald i Power (1993), McKenzie (2001), Mills (1993), Panas (2001), Peters (1992), entre d'altres.

Així, segons l'anàlisi R/S i els models ARFIMA no hi ha dependència a llarg termini en els rendiments dels actius del mercat anglès (Mills, 1993), resultat recolzat per altres estudis (Cheung, 1995). Aquest darrer, troba dependència tan sols per al mercat espanyol si s'aplica l'anàlisi R/S modificada, mentre que si s'apliquen models ARFIMA la dependència engloba a més a més als mercats austríac, italià i japonès. Sorprenent resulta que en un estudi d'un any posterior (Blasco i Santamaría, 1996) es conclogui que hi ha poca evidència de memòria a llarg termini utilitzant el mateix estadístic de l'anàlisi R/S modificada, tot i que l'evidència augmenta si el que s'analitza és el paràmetre d'integració en els models ARFIMA. Finalment, un dels estudis recents en la matèria (Henry, 2002) utilitzant també models ARFIMA detecta memòria a llarg termini per als mercats de Japó, Alemanya, Corea del Sud i Taiwan, mentre que no n'hi troba en els mercats del Regne Unit, Estats Units, Hong Kong, Singapur i Austràlia.

Potser en allò que es coincideix més en relació a l'anàlisi de la dependència en els mercats financers és que en cas de que n'hi hagi, és positiva a curt termini però negativa en el llarg. Tal com vàrem comentar en l'apartat anterior (pàg. 84), la tendència de l'exponent de Hurst és negativa respecte el temps, de manera que si bé per a períodes curts és $H \in (0'5,1)$ i la correlació és positiva, això no es contradiu amb un valor $H \in (0,0'5)$ per a intervals temporals més grans. Aquesta situació es descriu en dos articles paradigmàtics:

> "(...) *els nostres resultats suggereixen que pertorbacions en les rendibilitats mostren una correlació serial positiva sobre curts períodes i correlació negativa sobre llargs intervals*"[33]

> "(...) *les rendibilitats dels actius exhibeixen autocorrelacions positives en horitzons a curt termini i autocorrelacions negatives en horitzons a llarg termini*"[34]

Es fa palesa, a rel de la varietat de resultats obtinguts a excepció d'aquest darrer, la necessitat de limitar l'abast d'aplicació de la memòria com a dependència entre observacions en una sèrie temporal a la pròpia sèrie estudiada, sense poder fer-la extensiva a l'àmbit més general de mercat financer.

[33] Poterba i Summers (1988:53). [Traduït al català]
[34] Cheung (1995:598). [Traduït al català]

Notem que això últim comportaria la confusió entre concepte i mesura, portant a la incoherència de trobar mercats amb i sense memòria a la vegada, en funció de l'índex representatiu i interval temporal escollit i de l'instrument d'anàlisi emprat.

3.6. Anàlisi de la memòria-dependència amb XNA.

L'anàlisi de la memòria-dependència en les sèries temporals té com a finalitat millorar la modelització d'un conjunt d'observacions, incorporant els efectes de la relació entre valors passats de la variable estudiada. La principal dificultat amb la que ens trobem si utilitzem els instruments propis de l'econometria per assolir aquest objectiu, és el d'haver de definir una determinada relació funcional entre les observacions retardades, quan la dinàmica pròpia de la sèrie temporal acostuma a ser desconeguda. Tot i que s'han desenvolupat models i mètodes d'estimació més complexos per tal de recollir també la dependència no lineal en les variables, els models tradicionals i que encara s'empren en l'anàlisi de la memòria a llarg termini són modelitzacions de la dependència lineal. A més, ens trobem amb l'inconvenient de que no tan sols l'estimació dels paràmetres que defineixen els models és difícil, sinó que també ho és la interpretació a nivell financer dels mateixos.

A partir de la dècada dels 90, el desenvolupament de les XNA ha permès disposar d'una nova metodologia per al tractament de les sèries temporals, essent el major nombre d'elles destinades a fins predictius. Així, s'ha dissenyat una varietat de XNA que alimentades i entrenades amb dades històriques d'una variable, i afegint el valor – actual o passat – d'altres de relacionades, han tractat d'obtenir com a resultat el millor ajust possible a la sèrie temporal, i en conseqüència i com a projecció, l'anticipació del comportament futur. L'àmbit econòmic no ha quedat aïllat d'aquesta tendència. A Rosales et al. (2003) es troba una relació de treballs sobre XNA aplicades a la predicció de sèries temporals de variables econòmiques i financeres. Són molts els estudis realitzats en aquest sentit[35], esforç que es deu a la flexibilitat de l'instrument i als resultats obtinguts. Nosaltres proposem la mateixa metodologia però amb una finalitat diferent, la de comparar la bondat de l'ajust a una sèrie financera quan s'empren únicament dades històriques – memòria-dependència – i quan s'afegeixen alguns elements de la memòria col·lectiva. Complementàriament, avaluarem la importància d'alguna de les variables *input* de la xarxa sobre el valor de l'*output*.

[35] Bitzer i Omlin (2000), Moshiri i Cameron (2000), Pérez-Rodríguez i Torra (2001), Sitte i Sitte (2000), White (1996), Yao i Tan (2000), en són alguns exemples.

Ja hem mencionat en el capítol anterior (pàg. 64) que les XNA presenten com a principal avantatge sobre altres mètodes alternatius d'anàlisi de la dependència entre valors d'una sèrie temporal la seva facilitat per a detectar relacions de caràcter no lineal, i la característica pròpia d'aquest instrument de l'aprenentatge mitjançant la presentació d'exemples, amb la posterior abstracció per a tractar amb dades no presentades a la xarxa fins aleshores.

Per a l'ajust de sèries temporals s'han usat, generalment, tres tipus de XNA: les *feedforward*, les *feedforward/feedback*, també anomenades recurrents, i els models *híbrids*, en els que es barregen algorismes genètics, sistemes experts, lògica borrosa, etc. No ens podem decidir per una classe o altra si no tenim present el problema a tractar.

El nostre objectiu prioritari no és la predicció, sinó analitzar la memòria-dependència i les relacions entre ella i la memòria col·lectiva. No obstant, emprem com a suport de l'anàlisi la noció d'*importància*[36], i dins la divisió entre importància causal i predictiva ens decantem, com justificarem posteriorment, per la última. L'objecte d'estudi ens determina el tipus d'informació amb la que haurem d'entrenar la xarxa. D'entrada, tractarem amb dades històriques de la variable financera estudiada únicament i exclusiva. Així aconseguim que la matriu de pesos emmagatzemi el coneixement degut solament a la memòria-dependència. Si a més a més incorporem altra informació, com ara dades macroeconòmiques, detalls comptables, expectatives donades a conèixer en el mercat, etc., estem considerant altres factors que cauen fora d'aquesta accepció i que hem englobat en la memòria col·lectiva. Les diferents matrius de pesos ens serviran per a comparar els efectes deguts a una i a l'altra.

Aquest raonament també ens ajuda a determinar el tipus concret de XNA a utilitzar. La nostra proposta és l'aplicació del *perceptró multicapa* amb algorisme d'aprenentatge *back-propagation*, la implementació del qual ha estat descrit en el capítol 2. Si bé és cert que els models híbrids, com a combinació de metodologies, augmenten la bondat dels resultats obtinguts ja que es beneficien de les avantatges individuals dels instruments que agrupen, no és necessari ampliar una anàlisi en el que es tractaran

[36] Ens hem trobat amb dificultats terminològiques per a referir-nos al pes, o influència, d'un *input* sobre l'*output* ,sense emprar la mateixa paraula que assigna en una XNA un valor a la força de la connexió entre dues neurones. Per evitar aquesta dualitat i seguint a Sarle (2000), utilitzarem el terme 'importància', tot i la seva imprecisió, per anomenar el grau en el que un *input* intervé en la formació del valor de l'*output*. No hem optat per la més usual de 'sensibilitat' perquè sol associar-se a l'anàlisi de sensibilitat, basat en derivades parcials i, com detallarem en el següent apartat, aquesta no és l'anàlisi més adequada per a la nostra finalitat.

dades passades, i per tant, conegudes amb certesa, i de les que es vol analitzar solament la seva relació vers els valors futurs. D'altra banda, entre les xarxes *feedforward* i les recurrents, pot semblar que hauríem d'escollir les darreres, vist el que es diu d'elles[37]:

> "*Les xarxes neurals recurrents són instruments apropiats per a modelar sistemes variants en el temps (per ex., mercats financers* (...)*) Els mapes estàtics* [referint-se a les xarxes neurals *feedforward* multicapa] són *apropiats per al reconeixement de patrons- on els patrons espacials són independents del temps.*"

Malgrat els mercats financers siguin sistemes oberts i variants en el temps, nosaltres hem optat per l'ús d'una XNA amb connexions *feedforward*. Dos motius principals justifiquen l'elecció. El primer d'ells és la naturalesa de les dades amb les que entrenem la xarxa. Són dades històriques, conegudes, i cal que el sistema les reconegui com a valors ordenats cronològicament perquè és precisament la seva dinàmica la que volem modelitzar. D'altra banda, les experiències de treballs previs amb similars objectius, han obtingut bons resultats de l'aplicació d'aquestes xarxes a l'ajust de sèries temporals. A continuació descrivim un exemple de l'ús de les *back-propagation* en l'estudi de la memòria-dependència als mercats financers.

3.7. Ús de les xarxes back-propagation en l'anàlisi de la memòria-dependència. Una aplicació.

L'exemple que exposem a continuació pretén mostrar la manera en que les XNA poden ser emprades amb la finalitat d'analitzar i comparar els efectes del passat que són recollits per la memòria-dependència i aquells que es deuen a altres esdeveniments relacionats amb el coneixement històric propi dels mercats financers.

En concret, aplicarem les xarxes *back-propagation* per a l'anàlisi de la relació entre rendibilitats mensuals passades de l'índex *Dow Jones Industrial Average* (DJIA) amb la rendibilitat mensual del següent mes, i compararem els resultats amb la mateixa anàlisi a la que introduirem com a dada addicional una variable que reculli si el mes de sortida correspon o no a un octubre. La importància d'aquesta nova variable en l'anàlisi és una mesura de la influència dels records que evoca el mes d'octubre sobre la rendibilitat de l'esmentat mes. La noció d'"importància" en una XNA accepta dues interpretacions diferents (Sarle, 2000). La *importància causal* tracta de mesurar la

[37] Bitzer, T. i C.W. Omlin (2000:8-9). [Traduït al català]

variació en l'*output* davant canvis en el valor de l'*input*. En el nostre cas, els *inputs* són rendibilitats històriques i una variable qualitativa que pren valor 1 per a l'octubre i 0 per a la resta de períodes. En cap cas podem manipular el valor que prenen les variables. En conseqüència, el nostre enfocament és des de la perspectiva de la *importància predictiva*, mesurada com l'increment en l'error global comès per la xarxa quan ometem un dels *inputs*, el vinculat amb la memòria col·lectiva i que recull el record del mes d'octubre.

Cal esmentar que sols en casos particulars és possible trobar un únic nombre que mesuri la importància d'una variable explicativa sobre la variable explicada. De fet, aquesta és una qüestió que no ha estat resolta i que afecta especialment al camp de les XNA pel que fa a la interpretació de la matriu de pesos. A la pràctica, i davant la impossibilitat de trobar una solució general, s'opta per aplicar la mesura que més s'ajusta a cada problema concret. Ni tan sols en el cas més senzill de models amb funcions de transferència lineals, la mesura és única, ja que utilitzar el producte de matrius o les derivades parcials de la funció sortida com a mètode per a quantificar la importància de cada variable *input* és incorrecte si les dades no estan normalitzades, o fins i tot, si estant-ho, aquestes estan correlacionades. A més a més, per al càlcul de les derivades parcials és necessari que les variables siguin contínues, quan podem trobar-nos amb dades discretes, i el seu valor dependrà de la posició del punt en el pla al que s'apliqui, conseqüentment el resultat obtingut també tindrà només validesa local.

Vistes les dificultats per a determinar la influència de cada variable d'entrada en l'*output* d'una xarxa, s'han desenvolupat mètodes específics que la quantifiquen a partir de les matrius de pesos, com ara els de Garson (1991), Milne (1995) o Gedeon (1997). Tot i amb aquestes aproximacions, continuem trobant problemes degut a les formes que prenen les funcions de transferència entre les diferents capes de la XNA.

Sortosament, en el nostre cas, les variables *input* no presenten correlació i aquest fet ens permet, juntament amb el nostre interès per la importància predictiva d'una variable pròpia de la memòria col·lectiva, emprar la variació en l'error comès per la xarxa com una mesura útil de la importància d'aquella variable. Addicionalment, també hem calculat la mesura proposada per Gedeon (1997) i que pretén extreure a partir de les matrius de pesos de la xarxa, la contribució de cada variable *input* en l'*output* del sistema. A continuació detallem el procés seguit.

En l'exemple hem emprat les dades de l'índex novaiorquès des del gener de 1945 fins a l'abril de 2004, formant 699 patrons de 12 variables – rendibilitats mensuals de l'índex en els 12 mesos anteriors al valor que cerquem de sortida – per tal d'entrenar la XNA.

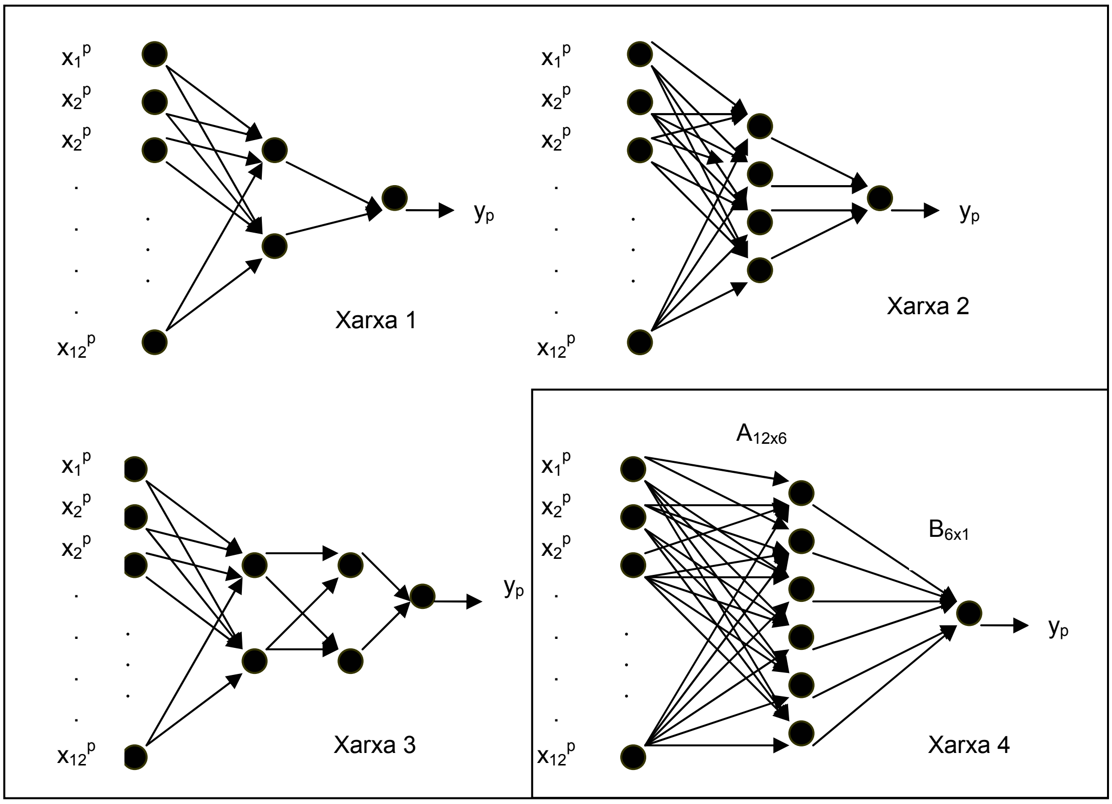
Una de les dificultats pràctiques de les xarxes *back-propagation*, ja mencionada en el capítol segon, és el de la indefinició de l'arquitectura més adient per tal d'emmagatzemar el coneixement dels patrons en els pesos del sistema. Aquesta mancança ha de ser coberta per l'encert en les decisions que prengui el dissenyador del sistema, convertint-se l'elaboració d'una XNA en tot un "art". El procediment més estès és el de provar diferents tipologies i optar per la que presenti el mínim error en l'ajust de les dades.

Donats els paràmetres de la xarxa (12 variables d'entrada, 1 de sortida i 699 patrons per a entrenar) i seguint les indicacions suggerides al capítol 2 per al nombre de neurones ocultes (pàg. 51), aquest hauria de trobar-se entre 4 i 11. Les diferents arquitectures provades i l'error comès, mesurat mitjançant l'error quadràtic mitjà[38] (MSE) es troben recollits en la taula 3.4. El gràfic 3.1. mostra esquemàticament l'estructura de cada XNA.

Xarxa	Nº capes ocultes	Neurones en les capes ocultes	MSE
Xarxa 1	1	2	0.0377707
Xarxa 2	1	4	0.0370248
Xarxa 3	2	2x2	0.0376394
Xarxa 4	1	6	0.0281102

Taula 3.4. Error comès en les xarxes *back-propagation* aplicades a l'estudi de la memòria-dependència.

[38] $MSE = \frac{\sum_{i=1}^{m} (output_i - y_i)^2}{M}$, on $output_i$ representa el valor real de la observació *i*, y_i és el valor de sortida de la xarxa per al patró *i*, i M és el nombre de patrons.



Gràfic 3.1. Arquitectures de xarxes *back-propagation* aplicades a l'estudi de la memòria-dependència.

Les XNA han estat entrenades amb els valors de la rendibilitat logarítmica mensual del DJIA a partir de la cotització al tancament en el darrer dia del mes[39], normalitzades dins l'interval tancat [-1, 1], emprant els 12 valors previs per a predir la següent observació. Com a suport informàtic hem utilitzat la *Toolbox* de Matlab de xarxes neurals artificials per a implementar les diferents tipologies, decidint-nos en tots els casos per funcions de transferència tangent sigmoidals, $f(x)=\frac{e^x - e^{-x}}{e^x + e^{-x}}$. Hem optat per aquesta funció perquè compleix una propietat necessària per a les xarxes *bac-propagation*, la diferenciabilitat, i a més a més, les seves imatges estan compreses entre els valors 1 i -1, a l'igual que la normalització a la que han estat sotmeses les dades que formen els patrons d'entrada.

Com a restricció per tal d'evitar bucles d'ineficiència, hem imposat que l'algorisme d'aprenentatge s'aturés quan la introducció de les dades a la xarxa assolís les 20.000 iteracions o bé quan l'error comès disminuís d'$1 \cdot 10^{-10}$. En totes les XNA analitzades,

[39] Les dades han estat obtingudes de *The Financial Forecast Center*, disponibles a la pàgina web: http://www.forecasts.org/data/data/djiaM.htm

l'algorisme s'ha aturat degut a la restricció del nombre d'iteracions, signe indicatiu que no s'ha trobat una matriu de pesos que s'ajusti acuradament a les dades.

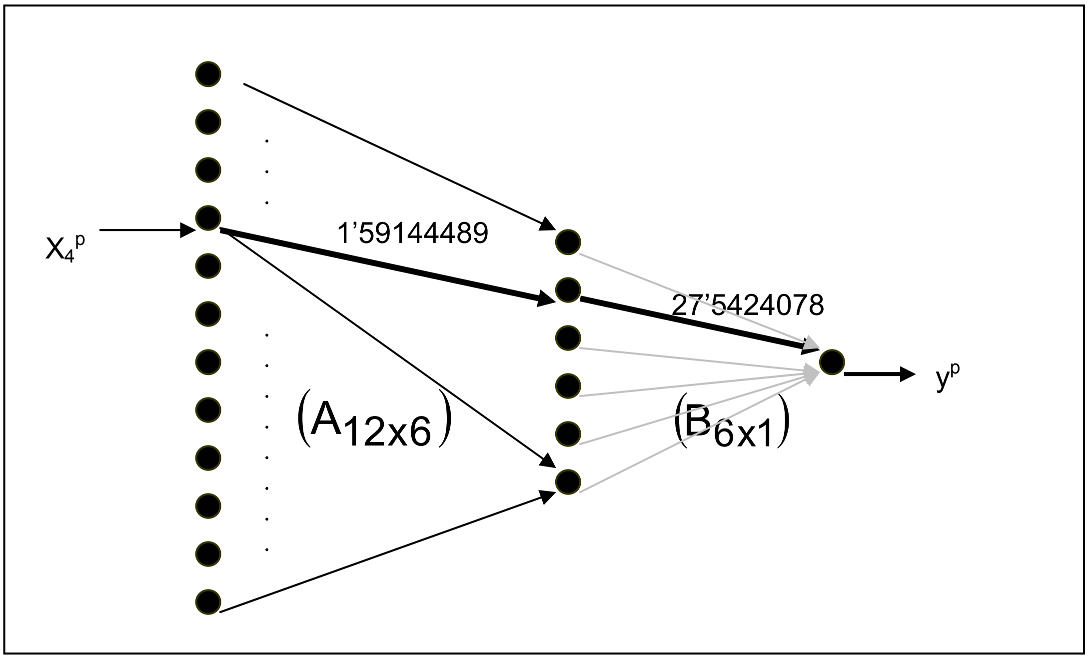
Tornant a la taula 3.4., observem que els errors han disminuït com a conseqüència d'augmentar el nombre de neurones en una única capa oculta, més que incrementant el nombre de capes ocultes. Malgrat ser elevats, podem considerar els errors assumibles pels nostres propòsits d'avaluar la influència d'incloure una nova variable en la xarxa, i per a mesurar la seva importància predictiva.

Seguirem l'anàlisi a partir de la xarxa 4, degut a que és la que presenta un error més petit. D'ella obtenim les matrius de pesos associades des de la capa d'entrada a l'oculta i des d'aquesta a la de sortida, anomenades A i B respectivament.

$$A_{12x6} = \begin{pmatrix} 0.69828713 & -0.72419494 & 0.31730268 & 224.668466 & 74.7406049 & 0.82285073 \\ -0.79600199 & 0.77627441 & -0.90633681 & 34.0197094 & 31.5274337 & 2.91542108 \\ -0.59705273 & 0.56156094 & -1.65480685 & 224.454157 & 16.3379394 & 4.2167137 \\ -1.59586603 & 1.59144489 & -0.12158057 & 172.572772 & -9.79154474 & -0.89334061 \\ -5.0865764 & 5.09529375 & -1.25175047 & 406.309706 & -16.8599596 & 0.96647308 \\ -2.48735333 & 2.51302223 & 0.27136738 & 215.301357 & -26.0827423 & 0.45866675 \\ 0.59128646 & -0.63643404 & 0.15926489 & 230.370653 & 2.54754449 & -0.78288449 \\ 3.24744336 & -3.27805207 & -0.28174098 & -229.119617 & 3.79208812 & 0.18622092 \\ 3.61946899 & -3.69586248 & 1.99130709 & 306.398148 & -63.9804124 & -3.02715986 \\ -1.23908507 & 1.24555512 & 1.34767683 & -0.57457175 & 45.3825542 & -2.35288367 \\ 3.74750858 & -3.73424032 & 0.56635885 & -268.374288 & -9.14198095 & 0.20551964 \\ 4.42460782 & -4.50609112 & -0.76154756 & -217.971417 & 59.3458593 & 1.221202 \end{pmatrix}$$

$$B_{6x1} = \begin{pmatrix} 27.8663787 \\ 27.5424078 \\ -47.4921448 \\ 0.16780744 \\ 0.4466233 \\ -1.17596092 \end{pmatrix}$$

On, per exemple, el pes 1'59144489 que es troba encerclat en la matriu A indica la força de la connexió entre la quarta neurona de la capa d'entrada (la que recull el valor normalitzat de la rendibilitat logarítmica del DJIA retardat 4 mesos) i la segona neurona de la capa oculta. En la matriu B, el pes 27'5424078, també dins un cercle, correspon a la sinapsi entre la segona neurona de la capa oculta i la de sortida de la xarxa. De forma esquemàtica, els representem en el gràfic 3.2.



Gràfic 3.2. Matrius de pesos en la xarxa 4.

Una mesura directa de la importància predictiva d'un *input* d'una XNA és la variació de l'error (MSE) en l'ajust en funció de la seva inclusió o no com a variable d'entrada. Degut a que hem entrenat les xarxes sense considerar el mes en que es produïa la rendibilitat, tornem ara a entrenar-la afegint una variable qualitativa que quantifiquem de la següent manera:

$$X^k_{13} = \begin{cases} 1 & \text{si k representa a un patró que recull la rendibilitat mensual d'un octubre} \\ -1 & \text{en altre cas} \end{cases}$$

Novament, emprem una xarxa *back-propagation* amb 6 neurones en una única capa oculta, una neurona en la sortida, però ara amb 13 neurones a la capa d'entrada. Mantenint les condicions d'aturada de l'algorisme d'aprenentatge, les corresponents matrius de pesos són ara:

$$C_{13x6} = \begin{pmatrix} -3.30079658 & 39.1892464 & 236.812885 & -1.98990547 & -0.41319007 & 5.60115646 \\ -7.98044676 & -19.2357583 & 201.801472 & -6.22473275 & 2.3512873 & -1.88099334 \\ 9.71831692 & 32.0098739 & 111.668583 & 6.38727986 & 2.66583927 & 3.25744863 \\ 1.02358965 & -9.23921582 & -363.16007 & -0.0183536 & 3.46695936 & -1.58083302 \\ -19.553153 & 6.09358209 & 561.422514 & -13.7131946 & -4.04735609 & 3.97758208 \\ -8.05731762 & 4.59041436 & -27.0557025 & -2.67761599 & 3.26796665 & -0.19713074 \\ -4.54964661 & -37.2729112 & -75.8519755 & -4.70297047 & 5.32857526 & -1.68744031 \\ 1.13602321 & -15.7909845 & 17.1511752 & -2.66623073 & 0.69069794 & -0.57581553 \\ 20.7602371 & -4.12217007 & -101.632257 & 17.6732899 & 4.51220473 & 1.90744415 \\ 10.6688022 & 0.44792413 & -146.998125 & 12.3149556 & 5.08556327 & -0.42452344 \\ 15.6001359 & -30.9242533 & -67.2615838 & 10.8895624 & 1.55322503 & -1.34774466 \\ -0.60221107 & -28.435065 & 36.5925136 & -2.60996155 & -0.27354318 & -4.5664958 \\ -5.23470312 & 0.37096415 & 57.4691852 & 11.2255112 & -7.66687439 & 4.96483981 \end{pmatrix}$$

$$D_{6x1} = \begin{pmatrix} 2.99122191 \\ -0.34705551 \\ -0.08458133 \\ -2.99286767 \\ -3.01145413 \\ 0.39953097 \end{pmatrix}$$

L'error comès per aquesta nova XNA és MSE=0.0285398. Si el comparem amb el corresponent a la xarxa 4 sense incloure la tretzena variable, MSE=0.0281102, s'observa que ha sofert un lleuger increment. Per tal de comprovar si l'augment en l'error és generalitzat, hem entrenat novament els quatre models de xarxa dissenyats per a analitzar la memòria-dependència. Els resultats obtinguts es resumeixen en la taula 3.5.

Model	MSE (sense variable octubre)	MSE (amb variable octubre)
Xarxa 1	0.0377707	0.0405086
Xarxa 2	0.0370248	0.0387584
Xarxa 3	0.0376394	0.0371376
Xarxa 4	0.0281102	0.0285398

Taula 3.5. Comparativa entre els MSE de les xarxes *back-propagation* aplicades a l'estudi de la memòria-dependència.

Quan incorporem en la XNA una variable que informa de si l'*output* correspon o no al mes d'octubre, els errors quadràtics mitjans augmenten, a excepció de la xarxa 3 on l'error presenta una petita disminució. Aquest fet és significatiu de que la nova variable no només no aporta informació rellevant per a l'*output*, sinó que és distractora i empitjora la bondat de l'ajust.

Abans de poder confirmar aquest resultat, però, cal tenir en compte que la mesura d'importància predictiva a partir de la variació del MSE és només vàlida per a models on els *inputs* no estiguin correlacionats, ja que en altre cas, no podríem separar quina part de l'error correspon al fet d'ometre una variable d'entrada i quina part és deguda als efectes derivats de la modificació dels pesos de les variables correlacionades amb l'omesa. El següent pas és, doncs, calcular la correlació entre les rendibilitats mensuals del DJIA. A partir de les dades que formen els 699 patrons d'entrenament de la xarxa, obtenim la matriu de correlacions següent:

$$\begin{pmatrix}
1 & 0.004316 & -0.025926 & 0.004081 & -0.000775 & 0.069973 & -0.046576 & 0.023809 & -0.025848 & 0.016357 & -0.011252 & -0.002764 \\
0.004316 & 1 & 0.004387 & -0.025998 & 0.005548 & -0.000937 & 0.071205 & -0.047105 & 0.024128 & -0.025292 & 0.016523 & -0.010376 \\
-0.025926 & 0.004387 & 1 & 0.004723 & -0.025567 & 0.006586 & -0.001387 & 0.070836 & -0.046948 & 0.025093 & -0.026003 & 0.017045 \\
0.004081 & -0.025998 & 0.004723 & 1 & 0.005506 & -0.024480 & 0.006374 & -0.001860 & 0.071047 & -0.045722 & 0.024353 & -0.025273 \\
-0.000775 & 0.005548 & -0.025567 & 0.005506 & 1 & 0.007825 & -0.027733 & 0.006765 & -0.002211 & 0.071851 & -0.047485 & 0.023574 \\
0.069973 & -0.000937 & 0.006586 & -0.024480 & 0.007825 & 1 & 0.007062 & -0.029125 & 0.007419 & 0.001330 & 0.069432 & -0.045209 \\
-0.046576 & 0.071205 & -0.001387 & 0.006374 & -0.027733 & 0.007062 & 1 & 0.008396 & -0.029885 & 0.005409 & 0.001748 & 0.067308 \\
0.023809 & -0.047105 & 0.070836 & -0.001860 & 0.006765 & -0.029125 & 0.08396 & 1 & 0.008409 & -0.030771 & 0.006452 & 0.001661 \\
-0.025848 & 0.024128 & -0.046948 & 0.071047 & -0.002211 & 0.007419 & -0.029885 & 0.008409 & 1 & 0.008751 & -0.031239 & 0.006385 \\
0.016357 & -0.025292 & 0.025093 & -0.045722 & 0.071851 & 0.001330 & 0.005409 & -0.030771 & 0.008751 & 1 & 0.006229 & -0.029859 \\
-0.011252 & 0.016523 & -0.026003 & 0.024353 & -0.047485 & 0.069432 & 0.001748 & 0.006452 & -0.031239 & 0.006229 & 1 & 0.004614 \\
-0.002764 & -0.010376 & 0.017045 & -0.025273 & 0.023574 & -0.045209 & 0.067308 & 0.001661 & 0.006385 & -0.029859 & 0.004614 & 1
\end{pmatrix}$$

De l'observació de la matriu, amb elements fora de la diagonal principal propers a zero, ja podem extreure que les dades no estan correlacionades i que, en conseqüència, l'argument sobre la mesura d'importància predictiva és correcte. Podem confirmar el resultat utilitzant, per exemple, el càlcul del determinant de la matriu de correlacions i el test d'esfericitat de Bartlett. Pel que fa al primer, quan més baix sigui el seu valor, més correlacionades estaran les variables, mentre que la incorrelació augmenta a mesura que el determinant s'apropa a la unitat. En aquest cas, el determinant és 0'933633. En relació al test de Bartlett, contrasta la hipòtesi nul·la de que la matriu és la identitat. Donat el valor de l'estadístic[40] (47,6) i els graus de llibertat[41] (66), el valor de la p del contrast és de 0'9574. Clarament, no es pot refusar la hipòtesi nul·la que comporta l'acceptació d'incorrelació entre les variables emprades per a entrenar la xarxa.

La interpretació que podem fer a rel dels resultats obtinguts és que en voler ajustar una sèrie temporal de rendibilitats mensuals, el fet de considerar el mes d'octubre com una dada rellevant per a l'anàlisi no suposa augmentar la qualitat de l'ajust. Ans el contrari, amb les noves matrius de pesos obtingudes degut a la inclusió d'aquesta nova variable, augmenta en conjunt l'error comès per la XNA.

[40] L'estadístic es calcula mitjançant la següent expressió: - [n-1-(2v+5)/6]·ln|M|, on n és el nombre de patrons, v és el nombre de variables i |M| és el determinant de la matriu de correlacions.
[41] Els graus de llibertat són $(v^2 – v)/2$, sent v el nombre de variables.

*

Per tal de completar el resultat anterior amb una mesura de la influència d'un *input* en el valor de l'*output,* que estigui dissenyada específicament per a les XNA, calcularem la proposada per Gedeon (1995). Segons aquest autor, si denotem per P_{jk} la contribució d'una neurona oculta *j* en la neurona de sortida *k*, calculada com:

$P_{jk} = \dfrac{|w_{jk}|}{\sum_{r=1}^{nh} |w_{rk}|}$, on w_{jk} és el pes de la connexió entre les unitats *j* i *k*, i n_h és el nombre de neurones ocultes de la xarxa; llavors, la contribució de la neurona input *i* en la neurona output *k* la denotem per Q_{ik} i s'obté de:

$Q_{ik} = \sum_{r=1}^{nh} (P_{ir} \times P_{rk})$, on P_{ir} és la contribució de la neurona d'entrada *i* en la neurona oculta *r*, definida en els mateixos termes que P_{jk}.

Aplicant aquest procediment a les dades del nostre exemple, obtenim:

$P_{1,1}$ =	0.3043970410	$P_{2,1}$ =	0.0353175636
$P_{3,1}$ =	0.0086072873	$P_{4,1}$ =	0.3045645192
$P_{5,1}$ =	0.3064559414	$P_{6,1}$ =	0.0406576471

i les contribucions de cada input en l'output són:

$Q_{1,1}$ =	0.0330796	$Q_{2,1}$ =	0.0664982
$Q_{3,1}$ =	0.0775972	$Q_{4,1}$ =	0.0336537
$Q_{5,1}$ =	0.1383093	$Q_{6,1}$ =	0.0567448
$Q_{7,1}$ =	0.0759567	$Q_{8,1}$ =	0.0202964
$Q_{9,1}$ =	0.1531963	$Q_{10,1}$ =	0.1092632
$Q_{11,1}$ =	0.0978374	$Q_{12,1}$ =	0.0226363
$Q_{13,1}$ =	0.1149303		

La contribució de la tretzena neurona d'entrada, la que recull la variable qualitativa per reflectir si la rendibilitat mensual del DJIA correspon o no al mes d'octubre, és la tercera en importància segons la mesura de Gedeon, després de la corresponent a la rendibilitat de nou períodes anteriors i a la de cinc mesos abans. Aquest fet sembla indicar que sí podria existir un lligam important entre la sortida de la xarxa i la variable que recull el mes d'octubre. Però si realitzem la mateixa anàlisi emprant els pesos de la xarxa 4 sense tenir en compte el mes del que es tracti, emprant únicament com a variables d'entrada les rendibilitats dels dotze mesos anteriors al que volem ajustar, s'obtenen els següents valors:

$P_{1,1}$ =	0.2661765	$P_{2,1}$ =	0.2630820
$P_{3,1}$ =	0.4536397	$P_{4,1}$ =	0.0016028
$P_{5,1}$ =	0.0042661	$P_{6,1}$ =	0.0112326

i les contribucions de cada *input* en l'*output* són:

$Q_{1,1}$ =	0.02981261	$Q_{2,1}$ =	0.05963368
$Q_{3,1}$ =	0.09176381	$Q_{4,1}$ =	0.03637262
$Q_{5,1}$ =	0.15541872	$Q_{6,1}$ =	0.06036277
$Q_{7,1}$ =	0.01966422	$Q_{8,1}$ =	0.07471552
$Q_{9,1}$ =	0.16516664	$Q_{10,1}$ =	0.08876087
$Q_{11,1}$ =	0.09718575	$Q_{12,1}$ =	0.12114279

On, novament, s'observa que les dues variables amb més influència en la formació de l'*output* són les rendibilitats de nou i de cinc mesos enrera. És a dir, encara que utilitzant la mesura de Gedeon s'obtingui un valor elevat de contribució de la variable octubre, comparant ambdues taules de contribucions, amb i sense la seva inclusió, ens adonarem que la diferència més significativa radica en la importància que se li atorga al darrer *input*, que recull la rendibilitat mensual de fa 12 mesos. La resta de neurones d'entrada varien poc la seva contribució a la neurona de sortida en un cas i altre. Si a aquest fet hi afegim que la variable amb més variació és precisament la que mesura la influència de la rendibilitat mensual d'un any abans, i per tant, ja recolliria els efectes dels octubres anteriors per al patró corresponent a aquest mes, podem concloure que la inclusió de la nova variable no afecta significativament la construcció de la sortida de la xarxa, i en tot cas, l'empitjora, ja que com hem vist anteriorment l'error mesurat pel MSE és més gran quan s'introdueix l'esmentada variable en l'anàlisi.

3.8. Consideracions finals.

En el present capítol hem posat de manifest les dificultats per a formalitzar la propietat de dependència entre observacions d'una mateixa variable separades cronològicament en una definició i tipologia úniques, així com la creixent complexitat en la modelització de sèries temporals amb memòria-dependència a llarg termini.

En aquest sentit, primer hem desglossat el ventall de definicions formals i criteris que ens permeten classificar una mateixa sèrie temporal en memòria a llarg termini, curt termini o sense memòria, segons diversos criteris, d'entre els que hem destacat la funció d'autocorrelació i la funció de densitat espectral. Cap criteri és superior a un altre, sinó que són concrecions formals diferents d'una mateixa característica. En segon lloc, hem fet un repàs dels models més emprats per a detectar memòria-dependència a llarg termini, explicitant l'anàlisi R/S, l'anàlisi R/S modificada i els models ARFIMA per la seva major aplicació en l'estudi de variables econòmiques i financeres. Novament, la dificultat en l'estimació dels paràmetres propis d'aquests models fa difícil etiquetar una sèrie temporal en funció del grau de record de les dades del passat.

A tot això hem d'afegir que els treballs empírics elaborats sobre la matèria tampoc han pogut aportar resultats aclaridors, ja que depenen en massa mesura tant de les dades concretes utilitzades com del model emprat.

Importants línies de recerca estan treballant per a millorar el tractament de la memòria-dependència i superar els inconvenients assenyalats en apartats anteriors. La majoria d'ells estan encaminats a l'objectiu de la predicció i s'allunyen de la nostra finalitat d'explicar com afecten els diferents tipus de memòria en el comportament d'un mercat financer. Per a tractar la memòria-dependència des d'una nova perspectiva, en la que es pugui partir inicialment d'una dinàmica desconeguda, on la dependència pugui ser no lineal, on els cicles no siguin periòdics, i on es pugui extreure coneixement de les dades històriques, i sobre tot, on es puguin comparar les diferents accepcions de memòria entre sí, la nostra aposta són les xarxes neurals artificials.

En particular, per a la memòria-dependència hem proposat el perceptró multicapa amb un algorisme d'aprenentatge *back-propagation*. L'adequació de la seva arquitectura per aproximar qualsevol funció ha estat la raó que ens ha dut a escollir-les.

De l'aplicació de les xarxes *back-propagation* a un cas concret en el que hem analitzat la importància de que sigui octubre en la variació de la rendibilitat mensual en el mercat financer americà, representat per l'evolució de la rendibilitat de l'índex *Dow Jones Industrial Average*, podem concloure que els resultats obtinguts recolzen els de les anàlisis estadístiques existents. És a dir, del nostre exemple també se'n deriva que l'"efecte octubre" és més psicològic que real ja que no produeix disminucions constants en la rendibilitat d'aquest mes. Al mateix resultat que pot arribar-se mitjançant estadístiques sobre el comportament de l'índex al llarg de la seva història, nosaltres hi hem arribat com a comparació dels errors comesos i de les matrius de pesos de les xarxes que hem aplicat en la seva anàlisi, amb l'avantatge que la metodologia de les XNA permet la comparació amb factors propis de la memòria col·lectiva.