

Documentos de trabajo – HISPALINK-Asturias

**TÉCNICAS DE REGRESIÓN FLEXIBLES
PARA LA MODELIZACIÓN Y
PREDICCIÓN DE LA ECONOMÍA
ASTURIANA**

DOCUMENTO DE TRABAJO 1/2002 (Noviembre 2002)

Han participado en la elaboración de este documento de trabajo:

Manuel Landajo Álvarez (landajo@correo.uniovi.es)
Rigoberto Pérez Suárez (rigo@correo.uniovi.es)
Departamento de Economía Aplicada
Universidad de Oviedo

D. Legal: AS-3708-2002

RESUMEN

En los últimos años se ha generalizado en el ámbito de la modelización y predicción económica el uso de una amplia gama de técnicas de regresión flexibles. Se trata de procedimientos de carácter no paramétrico o semiparamétrico, que se basan en ciertas familias de modelos no lineales. En este trabajo se analizan los rasgos principales de estas herramientas y sus conexiones con la modelización tradicional, comentando sus principales características técnicas y las posibilidades que ofrecen desde el punto de vista práctico en el ámbito de la predicción, haciendo especial énfasis en el ámbito de la economía asturiana.

Palabras clave

Predicción económica, modelos no lineales, métodos no paramétricos, economía asturiana.

ABSTRACT

In the latest years a number of flexible regression tools have emerged in the fields of economic modelling and forecasting. These are essentially nonparametric estimation devices, most commonly based on various classes of nonlinear models. In this paper, the most important among these techniques are reviewed and their connections to mainstream modelling and forecasting tools analysed. After a brief review of the main technical characteristics of these techniques we focus on the possibilities that they may offer to empirical forecasting, and specifically to economic forecasting in Asturias.

Keywords

Economic Forecasting, Nonlinear Modelling, Nonparametric Methods, Asturian Economic Forecasting.

1. INTRODUCCIÓN

En las dos últimas décadas ha ido creciendo entre los econométricos teóricos y los usuarios de las técnicas estadísticas de modelización y predicción económica la conciencia del carácter inherentemente aproximado de los modelos utilizados en la investigación empírica. Las características de los sistemas económicos - y las propias limitaciones informativas del investigador- obligan a considerar a las especificaciones paramétricas utilizadas en la modelización aplicada como meras aproximaciones, más o menos toscas, del verdadero proceso generador de los datos (P.G.D.) subyacente. A la vez que este reconocimiento, se ha ido ampliando el conocimiento teórico de los riesgos inferenciales que conlleva la especificación incorrecta de aspectos estadísticamente relevantes de los modelos, y en particular sus efectos sobre las estimaciones y predicciones, así como sobre la correcta interpretación de los contrastes de hipótesis.

Los esfuerzos de los investigadores teóricos han fructificado en forma de dos tipos de desarrollos: Por un lado se ha creado una *amplia teoría inferencial para modelos paramétricos incorrectamente especificados*, en la que se analizan los problemas de estimación y el tipo de hipótesis que pueden ser legítimamente contrastadas a partir de modelos incorrectos en alguno de sus aspectos. Por otra parte, y de forma paralela, se han ido desarrollado un conjunto de *técnicas flexibles de modelización y predicción*. Estas últimas son procedimientos de carácter no paramétrico, que no requieren para su funcionamiento la especificación previa de un modelo paramétrico del P.G.D. Dentro de las herramientas flexibles -a menudo de carácter multidisciplinar en sus orígenes, desarrollo y utilidades- se incluyen, junto a métodos más clásicos como los *estimadores de núcleo*, una amplia variedad de *estimadores de serie*, así como ciertas familias de

modelos no lineales, tales como las denominadas *modelos neuronales* o de *regresión conexionista*, además de un amplio conjunto de mecanismos de predicción basados en *sistemas de reglas*. Para todas estas herramientas se está desarrollando en la actualidad una teoría inferencial, que estudia sus capacidades desde el punto de vista de la inferencia no paramétrica.

En este trabajo se pretende en primer lugar hacer una presentación de los aspectos más relevantes de las técnicas de modelización flexibles. Los contenidos de esta materia se hallan repartidos en una voluminosa literatura de carácter multidisciplinar y considerable complejidad técnica. Por ello, pensamos que el objetivo prioritario —y el enfoque más útil del tema— ha de ser proporcionar una visión de conjunto de dicha materia, suficientemente general pero accesible al lector con intereses prácticos (las referencias al final del texto están seleccionadas para permitir una mayor profundización en los aspectos teóricos al lector interesado). Se pretende también que el texto tenga un carácter integrador: Las diversas técnicas son presentadas desde una perspectiva econométrica¹, al objeto de resaltar la coherencia y la compatibilidad de las mismas con los principios metodológicos de la modelización y predicción económicas. Este enfoque facilita también los análisis comparativos, permitiendo clarificar mejor las ventajas y limitaciones de los métodos flexibles frente a las técnicas paramétricas tradicionales.

El esquema del texto es el siguiente: En la sección 2 se analizan los problemas de especificación de modelos económicos empíricos y las ideas generales de la teoría inferencial para modelos incorrectamente especificados. En la sección 3 se pasa revista a las principales técnicas flexibles. La sección 4 se ocupa de algunos aspectos teóricos de carácter general relativos a los métodos no paramétricos y se comentan sus ventajas y limitaciones. En la sección 5 se analizan, desde el punto de vista empírico, algunos aspectos del uso de estas técnicas, haciendo referencia a algunas iniciativas que en la actualidad se están llevando a cabo dentro del equipo HISPALINK-Asturias al objeto de incorporar de forma efectiva estas técnicas a la mejora de las predicciones de la economía asturiana. Finalmente, en la sección 6 se presentan algunas conclusiones.

2. LOS PROBLEMAS DE LA ESPECIFICACIÓN DE MODELOS EN ECONOMÍA

Con frecuencia se suele señalar la considerable distancia de planteamientos existente entre la modelización económica —en su faceta del análisis teórico— y la construcción de modelos como tarea orientada a la predicción económica. La teoría económica se compone en una gran parte de un conjunto de modelos de carácter muy general, a menudo meramente cualitativo. Se trata de estructuras diseñadas para el análisis abstracto, con el objetivo de representar y ayudar a clarificar determinados aspectos del comportamiento de los sistemas económicos. Para dicha tarea no suele ser necesaria —ni conveniente— la especificación precisa de modelos cuantitativos con una estructura paramétrica prefijada. Por otra parte, a menudo la teoría sugiere que existen buenas

¹ Uno de los problemas prácticos que suele tener la audiencia económica (y también la de cultura estadística) para familiarizarse con las técnicas flexibles es su léxico, que a menudo proviene de campos tecnológicos. Sólo recientemente se ha comenzado a presentar estos métodos bajo un formato más adecuado a la condición de sus usuarios mayoritarios, que se encuentran precisamente en el ámbito de la economía, la empresa o la estadística.

razones para postular el carácter intrínsecamente no lineal de una gran variedad de relaciones económicas (Day, 1994).

Frente a esto, la modelización empírica parte de presupuestos bastante diferentes: sus prioridades son de carácter predictivo, y las limitaciones en la información disponible aconsejan rebajar las aspiraciones del investigador/usuario, recurriendo a especificaciones paramétricas precisas de los modelos, que sean tan simples como resulte posible y se alejen lo mínimo de la hipótesis de linealidad en las relaciones. Este conjunto de restricciones a la tarea del modelizador práctico, así como la distancia existente entre lo que permite la teoría y lo que aconseja el uso, llevan a plantearse las siguientes preguntas: (1) qué efectos tiene dicha simplificación, y (2) si este problema puede llegar a desaparecer con el tiempo, a medida que avance la teoría económica y consiga aportar condiciones de contorno más precisas para los modelos, y a medida que aumenten las dimensiones de nuestras bases de datos.

El econométra H. White ha sido uno de los autores que han tratado con mayor profundidad el problema de la especificación incorrecta de modelos y sus efectos, analizando ante qué clase de simplificaciones y errores de especificación son robustos los métodos de la inferencia estadística usados en modelización y predicción económica. Frente a la visión optimista, confiada en la posibilidad de alcanzar en el conocimiento de los sistemas económicos un grado de precisión similar al que se ha alcanzado en las ciencias físicas², autores como White (1994, *Capítulo I*) y Granger (1990, *Introducción General*) dan una respuesta mucho más pesimista a la cuestión de si el progreso de la ciencia económica puede aliviar el problema de la especificación incorrecta de los modelos. El primero de los autores señala lo siguiente:

"Debido a la extrema complejidad de la conducta económica, a la extrema dificultad de medir o incluso de definir adecuadamente los aspectos relevantes de los fenómenos económicos, y a que el economista suele tener poco o ningún control sobre los fenómenos económicos que estudia, la teoría económica está de raíz inherentemente limitada respecto del grado en el cual puede describir la realidad económica o formular juicios legítimamente contrastables (i.e. falsables) sobre la realidad económica ... Debido a que el economista empírico debe tratar con la realidad en toda su complejidad, es optimista en extremo esperar o creer que los modelos paramétricos habituales, económicos o probabilísticos, sean suficientemente adecuados para capturar esta complejidad." (White, 1994, pag. 2)³

² La idea tradicional del avance relativo de las ciencias físicas respecto a la modestia de los progresos de la economía parece excesivamente simplista. Ha de tenerse en cuenta que la mayoría de los sistemas económicos (p.e., los mercados financieros) son al menos tan complejos en su conducta como los más intrincados sistemas de las ciencias físicas. De hecho, Los físicos suelen denominar sistemas complejos a aquellos en los que interaccionan más de tres partículas u objetos. A partir de estas dimensiones desaparece la simplicidad, elegancia y precisión de los modelos, y muchos problemas se vuelven rápidamente inabordables.

³ Si, como sostienen muchos teóricos del análisis de sistemas, la complejidad de los sistemas es incompatible con la posibilidad de hacer afirmaciones precisas sobre los mismos (Zadeh, 1973), las perspectivas para el conocimiento económico y la predicción en dicho ámbito serían más bien dudosas. Sin embargo, lo cierto es que en muchas ocasiones nuestras técnicas proporcionan predicciones y diagnósticos razonablemente precisos, lo cual nos da cierto amparo frente a augurios pesimistas.

En los últimos años se ha desarrollado una amplia y sofisticada teoría⁴ (e.g., White 1981, 1982, 1994; Domowitz y White, 1982; Pötscher y Prucha, 1991a-b; Wooldridge, 1994), cuyo objeto es el análisis de los efectos estadístico-inferenciales que conlleva la especificación incorrecta de modelos econométricos, y que incluye también el análisis de otros tópicos como la robustez de los métodos ante una diversidad de problemas (ruido impulsivo, autocorrelación). En esencia, la teoría inferencial para modelos incorrectamente especificados garantiza que, bajo condiciones generales, los procesos de estimación en estos modelos no convergen hacia los verdaderos parámetros del sistema económico en estudio, sino a los de una aproximación óptima del mismo dentro de la clase paramétrica elegida.⁵ En su estado actual, la teoría incluye también el estudio de qué aspectos se pueden estimar y que tipo de hipótesis se pueden contrastar legítimamente a partir de modelos incorrectamente especificados en algún aspecto relevante (véase White, 1994). En la sección 5 analizaremos con más detalle estas cuestiones.

3. LA MODELIZACIÓN CLÁSICA. OBJETIVO E HIPÓTESIS BÁSICAS

La metodología estándar en el ámbito de la econometría y el análisis de series temporales dispone en la actualidad de un fundamento de carácter probabilístico, que históricamente ha ido sustituyendo el conjunto de técnicas de modelización de base empírica por modelos más sofisticados construidos sobre conjuntos de hipótesis estadísticas cuidadosamente formuladas. Los trabajos pioneros -entre muchos otros- de T. Haavelmo (1944) en el ámbito econométrico y de M. Kendall (véase Kendall y Ord, 1990, Harvey, 1989) en el ámbito de las series temporales están en la base de estos planteamientos. Por otra parte, en la actualidad la mayoría de la enorme variedad de métodos de uso común en la econometría y el análisis de series temporales son de carácter paramétrico, y se apoyan en las hipótesis de linealidad (en media) y especificación correcta del modelo.

El procedimiento habitual de modelización en el caso de la econometría comienza por asumir un P.G.D. de naturaleza estocástica. La información de que se dispone es una muestra compuesta por un número finito de observaciones extraídas de la población estudiada, es decir, un conjunto de pares (Y_t, Z_t) , $(t = 1, 2, \dots, n)$, donde Y_t y Z_t son vectores aleatorios (Z_t puede incluir tanto regresores estocásticos como no estocásticos, procesos de error y valores de Y retardados). Además, $Y_t = f_t(Z_t)$, donde f_t es una estructura, posiblemente cambiante en el tiempo (véase Domowitz y White, 1984). Habitualmente dicha estructura se reescribe como $f_t(\cdot, \theta)$, para reflejar el hecho de que es un elemento de entre una familia de funciones caracterizada por un parámetro θ que toma valores en un espacio paramétrico de dimensión finita Θ . De forma abreviada, $f_t(\cdot, \theta)$ es el modelo paramétrico asociado al P.G.D. En este planteamiento estándar sólo θ es desconocido, por lo que se constituye en el objeto básico de los procesos

⁴ Los primeros desarrollos corresponden principalmente a Huber (1964, 1967), que se planteó el problema de la especificación incorrecta en relación con el método de la máxima verosimilitud y propuso los M-estimadores y los primeros métodos de estimación robustos.

⁵ Un ejemplo concreto permite aclarar mejor este aspecto: si en un problema de regresión el investigador especifica erróneamente una superficie de regresión lineal mientras que la función $E(Y/X)$ es no lineal, los estimadores m.c.o., bajo condiciones generales, convergen estocásticamente hacia el modelo lineal más próximo (en el sentido mínimo-cuadrático) a la verdadera superficie de regresión.

inferenciales. Los objetivos del investigador serán la estimación de θ , el contraste de determinadas hipótesis relativas a dicho parámetro, o la predicción de determinadas magnitudes cuya conducta está determinada por el modelo. Todos estos aspectos deben llevarse a cabo sobre la base de la información muestral.

En el caso de las series temporales, la herramienta estándar, especialmente en predicción a nivel empírico, sigue siendo la modelización ARIMA, en la metodología de Box y Jenkins (1976). Se trata de un enfoque esencialmente lineal, aunque con un grado de flexibilidad elevado. La hipótesis clave es la posibilidad de reducir, mediante una transformación adecuada, la serie temporal a una nueva serie estacionaria en covarianza. Tras esto, la aplicación del Teorema de Descomposición de Wold (véase Doob, 1990, pag. 576) permite obtener una representación ARMA para la serie transformada (propriadamente el análisis no se refiere a la serie sino al proceso estocástico que se supone la ha generado). Las características relevantes del P.G.D. deben inferirse de una sola realización del proceso, y la ergodicidad del mismo es la hipótesis básica a tal efecto.

El esquema anterior puede generalizarse considerablemente, con la inclusión de variables exógenas, lo que da lugar a la clase de modelos ARMAX (*autoregressive moving-average with exogenous inputs*), y la consideración de modelos vectoriales como los VAR. Otros tópicos de reciente actualidad como los relativos a la cointegración y en general la regresión con variables no estacionarias no se han apartado esencialmente de las hipótesis de linealidad y especificación correcta del P.G.D. La inclusión de esquemas ARCH y otros similares, aunque aumenta la flexibilidad de los modelos (permitiendo ciertos tipos de no linealidades en varianza), sigue dentro del marco paramétrico, en el sentido de requerir la especificación previa de un modelo. Del mismo modo, los modelos estructurales de series temporales (véase Harvey, 1987), cuya principal fuerza se halla en su interpretabilidad y gran variedad de posibles especificaciones, son paramétricos, y generalmente lineales.

Una moderada separación respecto de los modelos lineales se encuentran en los modelos lineales por tramos y en sus variantes con transición suave entre regímenes. En el ámbito de las series temporales estos modelos han sido analizados por Tong (1990), Teräsvirta *et al.* (1992, 1994) y Teräsvirta (1994). Incluyen los TAR (threshold autoregression), STAR (smooth transition auto-regressive), y variantes como los modelos ESTAR. Una vez más, los modelos anteriores son de tipo paramétrico, requiriendo una fase previa de especificación.⁶

4. TÉCNICAS FLEXIBLES

4.1 Objetivos e ideas básicas

La idea básica de las técnicas flexibles, denominación que es sinónima de *no paramétricas*, es que sean los propios datos los que proporcionen toda la información al investigador, en el sentido de que no sea necesaria la especificación previa de un

⁶ En la literatura de series temporales se han propuesto algunas otras clases de modelos, como los *modelos bilineales* o los de *medias móviles no lineales*, aunque sus aplicaciones han sido escasas hasta la fecha, y los análisis teóricos de los mismos son más bien limitados. Teräsvirta *et al.* (1994) consideran más interesante el uso de otras clases de modelos más estructurados, como los STAR o las redes neuronales.

modelo, con lo que se eluden los riesgos que la misma conlleva. De manera más formal, en los procedimientos flexibles no se especifica un modelo paramétrico prefijado, sino que se va construyendo una sucesión creciente de modelos paramétricos, cuya complejidad se hace variar con el tamaño de muestra. Dicha sucesión debe construirse, como veremos más abajo, basándose en una familia “flexible” de modelos, de tal modo que se pueda obtener una aproximación cada vez más fina del fenómeno estudiado. El requisito básico que debe cumplir la sucesión de estimadores es el de consistencia, es decir, que converja en algún sentido relevante (aunque estocástico) hacia el objeto que se desea estimar (en general una superficie de regresión), de manera que asintóticamente se evite el efecto de la especificación incorrecta. Las técnicas flexibles, no paramétricas y semiparamétricas⁷, han encontrado utilidad en una variedad de problemas de predicción que incluyen -junto al análisis de regresión habitual- otros problemas de predicción como la *clasificación* y el reconocimiento de patrones. En la actualidad las capacidades de estimación flexible de estas técnicas están siendo utilizadas también para el desarrollo de una teoría del contraste de hipótesis de carácter netamente no paramétrico.⁸



Figura 1: Técnicas de modelización.

⁷ Los métodos semiparamétricos -o semi-no paramétricos- son realmente un caso particular de los no paramétricos. Elbadawi *et al.* (1983) acuñaron esta denominación para referirse a clases de modelos como la *forma de Fourier flexible*, que poseen una estructura de modelos lineales aumentados, en el sentido de que se construyen como suma de dos partes, una lineal o “paramétrica”, y otra “no paramétrica” que aporta la flexibilidad a los modelos. Por tanto, no existe pérdida de generalidad al hablar de métodos no paramétricos para referirnos a los métodos flexibles de forma inclusiva.

⁸ El objetivo de estas técnicas es contrastar hipótesis sobre características de la población (relevancia de variables, elasticidades, etcétera), pero sin recurrir a la necesidad de especificación previa de un modelo. La teoría es, en todo caso, muy reciente y se halla aún lejos de su madurez. Hasta el momento presente se han utilizado estimadores de serie (Hong y White, 1995) y de núcleo (Chen y Fan, 1999) para la obtención de contrastes de especificación. En un trabajo reciente, White y Hong (1999) han generalizado los M-tests al caso de los espacios paramétricos de dimensión infinita.

4.2 Clases de técnicas no paramétricas

4.2.1 Modelos lineales: Estimadores de serie

Estimadores de serie clásicos

Una parte importante de los métodos no paramétricos están basados en clases de estimadores lineales. Se basan en estructuras del tipo siguiente:

$$f(x) = \sum_{j=1}^m \beta_j \varphi_j(x)$$

Se trata, por consiguiente, de modelos generados mediante combinaciones lineales de cierta familia de funciones básicas $\{\varphi_j\}$. La condición que deben cumplir dichas combinaciones lineales es la de ser capaces de aproximar con precisión arbitraria las superficies de regresión, o en general los objetos que se desea estimar.⁹ Dos casos clásicos que permiten ilustrar la idea anterior son los polinomios algebraicos y trigonométricos, que corresponden, respectivamente, a los estimadores de serie construidos a partir de las funciones básicas de la forma $\varphi_j(x) = x^{j-1}$, $j=1,2,\dots$, y $\varphi_j(x) = \cos(b_j x)$ o $\varphi_j(x) = \text{sen}(b_j x)$. Es importante observar que las funciones básicas son prefijadas, en el sentido de que vienen dadas por la teoría matemática que sustenta los modelos, y por tanto no hay en ellas parámetros libres que deban ser estimados. Una vez determinado el número de términos del modelo, sólo hay que estimar los coeficientes β_j , para lo que basta utilizar algún algoritmo de mínimos cuadrados lineales. Esto dota a los estimadores de serie de una gran simplicidad computacional, y evita el riesgo de los óptimos locales asociados a la estimación no lineal. Además, el carácter lineal de los estimadores de serie facilita su estudio inferencial.

Los estimadores de serie han sido objeto de gran interés en el ámbito de la econometría tras el trabajo pionero de Elbadawi *et al.* (1983) en el que se proponía la forma de Fourier flexible (FFF) -una clase híbrida de estimadores de serie relacionada tanto con los polinomios algebraicos como con los trigonométricos- y se demostraba su utilidad para la estimación no paramétrica de elasticidades en sistemas de ecuaciones de demanda. Andrews (1991) ha completado dichos resultados, extendiéndolos a la estimación de funcionales econométricos lineales (derivadas, medias, integrales) de las densidades, y demostrando - en el caso de los datos transversales- la normalidad de la distribución asintótica de estos estimadores.

Estimadores de serie basados en B-splines y Wavelets

Recientemente se han incorporado al arsenal de los estimadores de serie otros modelos basados en funciones básicas diferentes de las algebraicas y trigonométricas. Es el caso de los B-splines y los wavelets. Estas nuevas clases de modelos se diferencian de los

⁹ En general el análisis se aplica a los espacios formados por funciones que admiten una representación

única del tipo siguiente $f(x) \equiv \sum_{j=1}^{\infty} \beta_j \varphi_j(x)$, donde (φ_j) es una familia prefijada de funciones básicas

(una base del espacio de funciones al que pertenece f , equipado en general con una distancia euclídea). Los estimadores se basan, por tanto, en desarrollos en serie truncados, en los que aparecen sólo m términos de la serie. A menudo las funciones básicas $\{\varphi_j\}$ se eligen ortogonales.

estimadores de serie anteriores en que sus funciones básicas φ_j , denominadas respectivamente B-splines y wavelets¹⁰, son funciones localizadas (i.e., se anulan fuera de un conjunto cerrado y acotado, o decrecen muy rápidamente a cero en el infinito).¹¹

Los B-splines son funciones básicas no negativas, polinomiales por tramos y generalmente de soporte acotado (i.e., nulas fuera de un cierto intervalo finito). El grado de los polinomios utilizados en los tramos determina el orden del B-spline (véase DeVore y Lorentz, 1993). Los estimadores de serie construidos a partir de B-splines son combinaciones lineales de un número finito de B-splines, i.e. estructuras del tipo

$$f(x) \equiv \sum_{j=1}^{\infty} \beta_j \varphi_j(x), \text{ donde } \{\varphi_j\} \text{ es una familia de B-splines, con orden y nudos prefijados por el usuario.}$$

Los wavelets son familias de funciones básicas generadas mediante traslaciones y dilataciones de una función originaria denominada *wavelet madre* la cual, a su vez, es también una combinación lineal de trasladados y dilatados de otra función, denominada *wavelet padre*. El resultado de este proceso da lugar a una familia de wavelets $\{\psi_j\}_{j \in J}$. Los estimadores de serie son (en el caso más simple) funciones de la forma genérica

$$f(x) = \sum_{j=1}^m \beta_j \psi_j(x).^{12}$$

Tanto los B-splines como los wavelets han sido estudiados extensamente en el ámbito de la teoría de la aproximación. Sólo más recientemente han empezado a explotarse sus capacidades para la construcción de estimadores no paramétricos y en otros problemas estadísticos. Chen y Shen (1998) han estudiado, en el contexto econométrico, las capacidades de estimación no paramétrica de wavelets, B-splines y otros estimadores de serie, obteniendo bajo condiciones generales, resultados de consistencia y normalidad asintótica. Su análisis, aún siendo considerablemente amplio, se limita por el momento al caso en que los datos se suponen generados por procesos estocásticos estacionarios y con dependencia débil. Algunas aplicaciones prácticas en problemas de econometría financiera han aparecido en Davidson *et al.* (1998) y Pan y Wang (1998). En el segundo de estos trabajos se proponen modelos de espacios de estados basados en wavelets.

¹⁰ Estos términos no se suelen traducir. B-spline significa "spline básico". El término wavelet (en francés *ondelette*) ha sido traducido -literalmente, aunque con escasa fortuna- como "olita".

¹¹ Los B-splines, a su vez, son un caso particular de los *splines*, que son funciones polinomiales por tramos, es decir, combinaciones lineales de funciones potenciales truncadas. Los B-splines tienen la ventaja sobre los splines genéricos de su localización (véase DeVore y Lorentz, 1993; Chui, 1988).

¹² Los wavelets son una *base incondicional universal* en una gran cantidad de espacios de funciones de interés práctico. Por ejemplo, toda función perteneciente al espacio de las funciones de cuadrado

integrable $L^2(\mathbb{R}) \equiv \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid \int_{\mathbb{R}} |f(x)|^2 dx < \infty \right\}$ admite una representación única del tipo

$f(x) = \sum_{j=1}^{\infty} \beta_j \psi_j(x)$, donde $\{\psi_j\}$ es una base de wavelets dicho espacio. Los estimadores de serie obtenidos a

partir de wavelets son por tanto series truncadas. Una exposición completa de la teoría matemática de los wavelets se halla en Meyer (1990).

4.3 Modelos no lineales

4.3.1 Estimadores de núcleo

El problema de la estimación no paramétrica de funciones de densidad -y en general de funcionales de las mismas- tiene una larga tradición en el ámbito de la estadística, aunque ha suscitado mayor interés a partir de los años ochenta. Una de las líneas clásicas, y que ha fructificado en forma de una amplia literatura, deriva de los trabajos pioneros de autores como Fix y Hodges (1951), Rosenblatt (1956) y Parzen (1962, 1979), y ha recibido la denominación genérica de estimación mediante núcleos (*kernel estimation*). La idea básica es la siguiente: para una variable aleatoria X con función de densidad $f(x)$, y dada una muestra aleatoria (x_1, x_2, \dots, x_n) de la misma, se estima $f(x)$ mediante la expresión:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

donde la función $K(\cdot)$ se denomina *núcleo*, y habitualmente se elige como tal una función de densidad, a la que se impone condiciones de continuidad, acotación y localización adecuadas (en general poco restrictivas), y $h > 0$ (conocido como tamaño de ventana) es un parámetro de alisado cuyo valor debe elegir el usuario. Una elección típica para el núcleo es la función gaussiana, basado en la densidad normal estándar, aunque existen otras muchas posibilidades (p.e., uniforme, Epanechnikov). Los estimadores de núcleo son esencialmente mecanismos de alisado, muy relacionados con el histograma. El planteamiento de los estimadores de núcleo se extiende sin dificultad al caso en que X es un vector aleatorio de dimensión finita.

A partir del estimador anterior se estiman analógicamente las superficies de regresión. Un caso clásico es el estimador de Nadaraya-Watson, que en el caso de núcleo gaussiano (supuesta X escalar), para una muestra aleatoria simple $(x_j, y_j)_{j=1,2,\dots,n}$, adopta la siguiente forma:

$$\hat{E}(Y/X=x) = \frac{\sum_{j=1}^n y_j \exp\left(-\frac{1}{2}\left(\frac{x - x_j}{h}\right)^2\right)}{\sum_{j=1}^n \exp\left(-\frac{1}{2}\left(\frac{x - x_j}{h}\right)^2\right)}$$

La expresión anterior se obtiene sustituyendo en la definición de $E(Y|X)$ la densidad $f(y|x)$ por el estimador de núcleo de la misma. De manera análoga se obtienen estimadores para volatilidades (varianzas) condicionales y otros funcionales de interés, como elasticidades y medidas de curvatura.

Bajo condiciones generales sobre $f(x)$, el núcleo K y la velocidad de decrecimiento del ancho de ventana h , se demuestra que los estimadores de núcleo son consistentes (p.e., en el sentido de la media cuadrática). Evidentemente, también son no lineales respecto del parámetro de alisado h .¹³

¹³ Para un análisis exhaustivo de los estimadores de núcleo, sus orígenes y el estado del arte con relación a los mismos puede consultarse Simonoff (1996).

Se han señalado como limitación de los estimadores de núcleo su ineficiencia (en el sentido estadístico del término) con relación a los métodos paramétricos, su sensibilidad al problema conocido como *el curso de la dimensionalidad* y su comportamiento de frontera. También se ha señalado la dificultad de determinar el valor del parámetro de alisado, incluso aunque se recurra a técnicas de validación cruzada, y su carácter de *cajas negras computacionales*, es decir, estructuras no interpretables. Por ello, los estimadores de núcleo se recomiendan a menudo para fines de análisis exploratorio, al objeto de ayudar a buscar un modelo paramétrico adecuado.

En el ámbito de la econometría, los estimadores de núcleo han sido estudiados ampliamente por Ullah (1988) y Teräsvirta *et al.* (1994). En el análisis de series temporales han sido aplicados a la estimación de volatilidades condicionales en mercados financieros por Gallant *et al.* (1993), además de ser métodos clásicos utilizados en el análisis espectral (véase Kendall y Ord, 1990).

4.3.2 Modelos neuronales

Las redes neuronales artificiales (RNA), también llamadas *modelos de regresión conexionistas*, son una clase de modelos cuyo origen histórico se halla en el ámbito de la biología,¹⁴ donde se introdujeron como representaciones simplificadas de los sistemas de neuronas naturales. Las aportaciones pioneras de A. Lapedes y R. Farber (1987) mostraron por primera vez la posibilidad de utilizar las RNA para la predicción de series temporales caóticas. Los resultados prácticos obtenidos en estas primeras aplicaciones permitieron resolver algunos problemas que no hallaban tratamiento satisfactorio en las técnicas tradicionales.¹⁵ Desde entonces se ha desarrollado una extensa literatura sobre aplicaciones de las redes neuronales en tareas de predicción económica y empresarial - especialmente amplia en el caso de la predicción de series financieras- así como en una gran variedad de problemas de clasificación (véase Cheng y Titterington, 1994). Más recientemente se han iniciado las aplicaciones con orientación específicamente econométrica, a partir de trabajos como los de Maasoumi (1994) y Swanson y White (1995a-b).

Desde una perspectiva estadística los modelos neuronales son simplemente mecanismos de regresión basados en formas funcionales flexibles, que puede ser utilizados como estimadores no paramétricos en problemas de modelización y predicción económica. Aunque existe una enorme diversidad de modelos y técnicas de estimación que se pueden ubicar dentro del paradigma neuronal, a los efectos que aquí nos ocupan dividiremos los modelos neuronales en las categorías de redes unidireccionales y redes recurrentes.

¹⁴ El origen histórico de los modelos neuronales ha determinado el léxico que hasta hace poco era de uso común en dicho ámbito, que venía heredado de las ciencias naturales y la computación.

¹⁵ La utilidad de las técnicas neuronales y de otras técnicas ligadas a la inteligencia artificial, así como el interés de su mayor incorporación en la economía y la empresa españolas, han sido recientemente señaladas por diversas recomendaciones de la Unión Europea, en las que se hace énfasis en la conveniencia de que nuestro país incremente la aplicación de estos métodos en todos los ámbitos (diario 'Cinco días', miércoles, 21 abril de 1999).

Las redes unidireccionales

Son combinaciones lineales finitas de funciones básicas, denominadas *funciones de activación neuronal* (o, simplemente, *no linealidades*). Una estructura de las más populares es la denominada *perceptrón multicapa*, que suele presentar la forma siguiente:

$$Y = \beta_0 + \sum_{j=1}^m \beta_j \Phi(\alpha_{0j} + \alpha_j \cdot X'),$$

o bien, en una versión algo más adecuada para fines prácticos, la forma de un modelo lineal aumentado:

$$Y = \gamma_0 + \gamma \cdot X' + \sum_{j=1}^m \beta_j \Phi(\alpha_{0j} + \alpha_j \cdot X'),$$

donde X es el vector de variables de entrada de la red, Y es la variable de salida y $\Phi(\cdot)$ es una función de transferencia no lineal, a la que habitualmente se exige que posea las características de una función de distribución acumulativa, es decir, que sea una función sigmoideal.

Existen otras muchas clases de RNA unidireccionales, tales como las *redes de base radial*, donde la función básica es una función de activación localizada, y que se hallan muy relacionadas con los estimadores de núcleo vistos más arriba.

En principio los modelos neuronales unidireccionales son aptos para su uso tanto en contextos de regresión con datos transversales como en el caso de las series temporales. En este último caso, cuando en el regresor X se incluyen valores retardados de la variable Y , los modelos neuronales dan lugar a estructuras autorregresivas no lineales o NAR (*non-linear auto-regressive*). Por otra, parte, los modelos neuronales del tipo perceptrón presentan una gran similitud con otra técnica estadística no paramétrica, la *regresión mediante búsqueda de proyecciones* o PPR (véase Friedman y Stuetzle, 1985), aunque esta última es algo más compleja en su implementación.

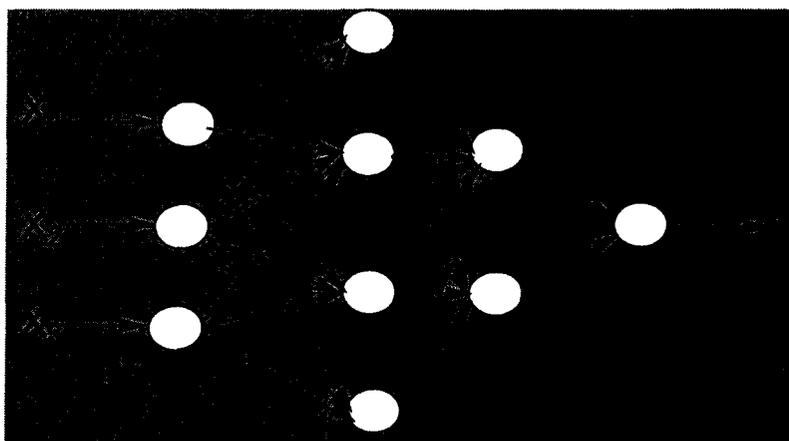


Figura 2: Representación esquemática de una RNA unidireccional.

Las redes recurrentes

Son una clase de modelos neuronales cuya rica estructura dinámica les hace especialmente adecuados para la modelización y predicción en el ámbito de las series temporales.

Desde una perspectiva econométrica, las redes recurrentes son una clase de modelos dinámicos con variables latentes. Así, una de las estructuras más típicas, la red Elman, es caracterizada por Kuan y White (1994a) del modo siguiente:

$$\begin{aligned} Y_t &= F_t(X_t, \theta) = \beta C_t && \text{(ecuación de medida)} \\ C_t &= G(X_t \gamma + C_{t-1} \delta) && \text{(ecuación de estado)} \end{aligned}$$

donde β , γ y δ son vectores de parámetros, C_t es un vector de variables de estado (denominado "contexto" en la jerga neuronal) y $G(\cdot)$ es una función de transferencia (a menudo de tipo sigmoïdal), X_t es la entrada del sistema e Y_t la salida. Este tipo de redes presenta una clara conexión con los modelos lineales de espacio de estados (véase Harvey, 1989), y en particular con una clase de modelos adaptativos propuesta por Priestley (1988) para el tratamiento de series temporales.

También es importante observar la conexión de las redes recurrentes con la modelización ARIMA. Desde este punto de vista, las estructuras recurrentes vistas pueden representarse como modelos ARMA no lineales (NARMA, o NARMAX si incluyen también variables exógenas), en los que junto a valores retardados de Y aparecen procesos de tipo MA. Una forma típica sería la siguiente:

$$Y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{-q}) = \sum_{j=1}^m \beta_j \sigma \left(\sum_{i=1}^p \alpha_{ij} y_{t-i} + \sum_{i=1}^q \gamma_{ij} \varepsilon_{t-i} \right) + \varepsilon_t$$

donde $\sigma(\cdot)$ es una función de transferencia sigmoïdal, y se imponen restricciones adicionales adecuadas sobre las propiedades estocásticas de los procesos $\{Y_t\}$ y $\{\varepsilon_t\}$.

Por tanto, en el contexto de las series temporales, las redes unidireccionales pueden considerarse como estructuras AR no lineales, mientras que las recurrentes se identifican con estructuras ARMA no lineales (Connor *et al.*, 1994; Cottrell *et al.*, 1995). En los modelos se pueden incluir también análisis de intervención y otros aspectos comunes en el análisis de series temporales.

La "economía neuronal"

Los trabajos de diversos autores han llevado a la aparición de un cuerpo de literatura especializada dedicado al estudio de los modelos neuronales desde la perspectiva econométrica. Este conjunto de contribuciones ha conseguido clarificar el papel de las RNA dentro del ámbito de la modelización y la predicción económica, proporcionando un conocimiento riguroso de las propiedades estadísticas de estos modelos, del que previamente se carecía, y permitiendo una mayor corrección en el uso en la práctica de los mismos. Los tópicos básicos de este conjunto de contribuciones se refieren al análisis de las capacidades de aproximación y aprendizaje o estimación de estos modelos. Respecto de este último, cabe citar entre las contribuciones de mayor interés, los trabajos teóricos de White (1987, 1989a-d, 1990, 1994), Gallant y White (1992),

Kuan y White (1994a-b), Xu et al. (1994) o Geman *et al.* (1992).¹⁶ Entre los tópicos de dicha literatura se incluyen la estimación analógica de estos modelos, su uso en contrastes como el de no linealidad, y el estudio de los algoritmos recursivos de estimación neuronal.

Los algoritmos de entrenamiento neuronal y la modelización dinámica en economía

Para la estimación (o entrenamiento) de los modelos neuronales - que suelen estar fuertemente parametrizados- resulta a menudo inadecuado el uso de las técnicas de optimización no lineal habituales en econometría. Por ello, se han desarrollado algoritmos específicos de estimación, más ligeros desde el punto de vista computacional. Se trata, en su mayoría, de procedimientos de carácter recursivo. El más clásico - aunque limitado en sus capacidades- es el algoritmo denominado BP, el cual es esencialmente una implementación recursiva de los mínimos cuadrados no lineales. Los algoritmos recursivos suelen dar estimadores menos eficientes que los analógicos, pero a menudo son más ágiles y simples. Diversos trabajos recientes han conectado estos algoritmos con algunos tópicos de gran importancia en el ámbito de la estadística, en particular con *la teoría de la aproximación estocástica, los filtros de Kalman extendidos y la teoría de los m-estimadores recursivos*.

Es interesante señalar una ramificación bastante sorprendente -y muy poco conocida- de los estudios neuronales. Se trata de la conexión matemática entre los modelos neuronales y la teoría económica. Apuntamos dos ejemplos: Por una parte, distintos trabajos llevados a cabo en los últimos años, relativos a la modelización del aprendizaje de los agentes económicos y su convergencia hacia el equilibrio correspondiente a la hipótesis de expectativas racionales (véase Marcet y Sargent, 1989), han conducido a idénticas clases de modelos, y al estudio de *los mismos* algoritmos recursivos de aprendizaje, que los que aparecen en la estimación de los modelos neuronales. Las aportaciones de Kuan y White (1994a-b) son una importante generalización de dicha literatura al contexto no lineal (y bajo procesos dependientes). Por otra parte, se ha señalado por diversos autores (p.e., Kosko, 1992; Johnson *et al.*, 1994) la equivalencia matemática entre los teoremas de equilibrio general en economía y los teoremas de equilibrio dinámico en sistemas neuronales. La explicación de este fenómeno se relaciona con que ambos tipos de resultados se refieren al equilibrio de sistemas dinámicos en los que interactúan un número elevado de agentes (agentes económicos, neuronas), cuyo efecto individual sobre el estado global del sistema es pequeño. Sin pretender entrar en detalles, resulta notable la simetría entre los comportamientos de sistemas en ámbitos tan distintos como la biología, la computación y la economía, que parecen plegarse a los mismos patrones matemáticos. No parece exagerado suponer que este hecho ofrece -en el medio plazo- posibilidades que hoy son difíciles de prever y abundantes trasvases metodológicos entre estos campos.

4.3.3 Sistemas de reglas

Finalmente, haremos una breve referencia a otro conjunto de técnicas que pueden incluirse dentro de los métodos no paramétricos. Nos referimos a las técnicas que se basan en la codificación del conocimiento en forma de sistemas de reglas. Se trata de mecanismos de uso común en problemas de modelización, predicción y control,

¹⁶ Una revisión de la literatura estadística de las RNA - adecuada como lectura introductoria- se encuentra en Cheng y Titterington (1994).

especialmente en problemas de clasificación, en ámbitos como las finanzas y la tecnología. De forma esquemática, la estructura de estos modelos es la siguiente: El espacio X se parte en una serie de regiones $\{A_1, A_2, \dots, A_m\}$, para cada una de las cuales se construye una regla, que indica a grandes rasgos cómo se comporta la salida del sistema (Y) en dicha región (véase figura 2):

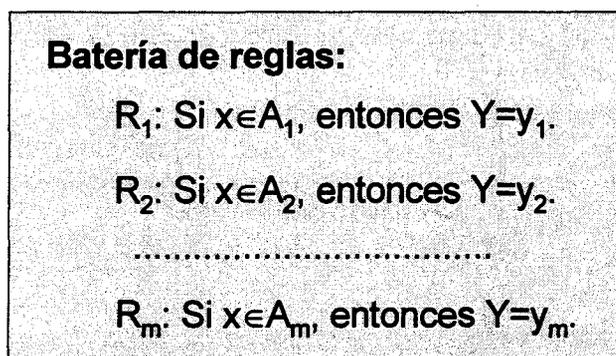


Figura 2: Representación esquemática de un sistema de reglas.

Aunque se suele pensar en los sistemas de reglas como sistemas cualitativos, lo cierto es que se pueden analizar sus propiedades numéricas del mismo modo que en los demás métodos flexibles. Así, por ejemplo, las capacidades de estimación no paramétrica de los sistemas de reglas denominados IA (abreviatura de *inteligencia artificial*), que matemáticamente se basan en funciones escalonadas, pueden analizarse aplicando resultados de Chen y Shen (1998). Las capacidades matemáticas de aproximación y el comportamiento estadístico de otra variante, los sistemas de reglas de lógica multivalente o *sistemas borrosos*, se pueden analizar mediante procedimientos análogos a los utilizados en el estudio de los modelos neuronales (Landajo, 1999).

5. PROCESOS DE APRENDIZAJE Y TEORÍA INFERENCIAL. UN ENFOQUE INTEGRADO

En esta sección analizaremos con más detalle los aspectos básicos de tipo inferencial que aparecen en relación con la modelización económica, y específicamente en relación con las técnicas flexibles. Salvo donde se indique expresamente, las ideas y conceptos son de validez general para todas las técnicas.

En el análisis teórico se deben considerar tres posibles escenarios: Los dos primeros son de carácter paramétrico (los modelos correctos y los modelos paramétricos incorrectos), y el tercero, que es el objetivo en sentido estricto de nuestro análisis, se refiere a la estimación no paramétrica. En principio, en los tres casos el objetivo es el mismo: la estimación de una función f , que puede ser la superficie de regresión $E(Y/X)$. Para la construcción de los modelos se recurre a una clase paramétrica, de la forma genérica $A_m(x, w)$, donde m es un indicador de la complejidad del modelo (p.e., el número de términos de un estimador de serie) y w un vector de parámetros libres, una vez prefijado m .

5.1 Métodos paramétricos

Caso 1: Corresponde a la teoría de la estimación clásica. La hipótesis fundamental es que el modelo está correctamente especificado: existe una elección m^* y un vector de parámetros tales que $E(Y/X) \equiv A_{m^*}(X, w^*)$.

Caso 2: Se refiere a la estimación paramétrica con modelos incorrectamente especificados: Se fija la complejidad m del modelo y el objetivo es estimar el vector w , con la particularidad de que no existe una elección de w^* para la cual el modelo coincida con la superficie de regresión. Como hemos comentado más arriba, el objeto que se estima es un modelo $A_m(X, w^*)$ que aproxime óptimamente a $E(Y/X)$, aunque ambas funciones no coincidirán.

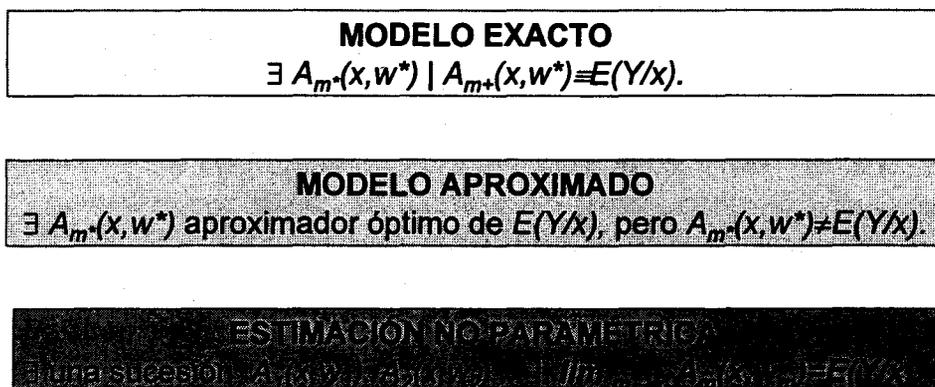


Figura 3: Aspectos inferenciales. Casos posibles.

El primer efecto de la especificación incorrecta de aspectos relevantes del modelo va a ser el hecho de que, utilizando un modelo incorrecto, se cometerán errores de predicción que podrían evitarse con la especificación correcta. Además, dichos errores no se podrán corregir por muy grandes que sean las muestras utilizadas para estimar los parámetros del modelo. Por otra parte, la especificación puede ser correcta respecto de ciertos aspectos e incorrecta respecto de otros. Este último es un fenómeno especialmente frecuente en los modelos dinámicos.

Una alternativa a los riesgos de la especificación incorrecta es el uso de métodos de estimación no paramétricos.

5.2 Métodos no paramétricos

En la estimación no paramétrica, la idea es que f puede no ser representable mediante un modelo paramétrico de complejidad finita $A_m(X, w)$. Pero si la clase de los modelos $A_m(X, w)$ es flexible, será posible construir un esquema de estimación no paramétrico para f a partir de la misma. El análisis de este proceso se puede basar en un procedimiento de carácter general conocido como *el método de los tamices*, debido a U. Grenander (1981).

La estimación mediante tamices

Se puede demostrar que las diferentes clases de estimadores flexibles vistas en la sección 5 se pueden hacer corresponder a diferentes particularizaciones de un mismo esquema general. La estimación de tamiz suministra dicho marco general, y facilita una

visión de conjunto de todas estas técnicas. A continuación analizaremos sucintamente las ideas básicas de dicho método.

Esencialmente, la construcción y el análisis de estimadores flexibles o no paramétricos se apoya en dos pilares básicos. En primer lugar, un conjunto de resultados matemáticos, conocidos como *teoremas de aproximación universal*, y en segundo lugar, la construcción de un mecanismo de *tamiz* propiamente dicho.

A) *Teoremas de Aproximación Universal*: La construcción de todo procedimiento de estimación no paramétrica requiere disponer de una familia de modelos flexibles. Dicha flexibilidad es una propiedad de carácter puramente matemático, que se plasma en un conjunto de teoremas de aproximación universal. La idea es que la función f que se desea estimar forma parte de un cierto espacio de funciones E , y para construir los modelos se debe recurrir a una clase aproximadora de funciones que denominaremos A . Los enunciados de los teoremas de aproximación universal adoptan siempre una misma forma genérica: *Dada una clase de funciones E (a la que pertenece f) y una clase de aproximadores A (la cual se utiliza para construir los modelos), se demuestra que cualquier elemento de E puede ser aproximado (en el sentido de una determinada métrica) con precisión arbitrariamente elevada por medio de elementos de A .*¹⁷

Las diversas clases de procedimientos vistas (series, redes neuronales, sistemas de reglas) tienen como soporte fundamental alguna clase flexible de funciones (polinomios, combinaciones lineales de wavelets, funciones escalonadas, etcétera), que forman el soporte "funcional" del método, y para la que existe algún tipo de teoremas de aproximación universal.

B) *El método de los tamices*: Un tamiz (*sieve*) es una sucesión creciente de modelos paramétricos, cuya complejidad se hace crecer en función del tamaño de muestra disponible. Si este procedimiento se lleva a cabo a partir de una familia de modelos flexibles, es decir, que verifique un teorema de aproximación universal adecuado, y si la complejidad de los modelos (i.e., el número de parámetros a estimar) crece ilimitadamente, aunque a una tasa suficientemente lenta como para evitar el sobreajuste, se puede obtener, bajo condiciones generales, que el tamiz converja estocásticamente hacia la superficie de regresión, o más generalmente, hacia el objeto que se desea estimar.

Las técnicas no paramétricas analizadas más arriba pueden considerarse casos particulares del método de los tamices: en todas ellas se utilizan familias de funciones flexibles y se construyen a partir de ellas una sucesión creciente de modelos que sea consistente para estimar la superficie de regresión f . Por tanto, *la idea de la estimación no paramétrica es que en el límite, i.e., cuando el tamaño de muestra y la complejidad de los modelos aproximadores sea suficientemente elevada, se produce la convergencia ("en probabilidad", "casi seguro" o "en media cuadrática") hacia la función que se desea estimar, sin necesidad de que el investigador haya especificado previamente una forma paramétrica para el modelo.*

En todos los casos vistos el método de los tamices proporciona un marco general para el proceso de estimación, facilitando además el análisis de las propiedades inferenciales de los estimadores no paramétricos. En los trabajos de Geman y Wang (1982), Shen

¹⁷ En otros términos, todo elemento de E es límite de una sucesión de elementos de A , o equivalentemente, A es denso en E .

(1997), Chen y Shen (1998), se analizan, desde el punto de vista de la estimación mediante tamices, una amplia variedad de estimadores no paramétricos, incluyendo aspectos como la consistencia y la normalidad y eficiencia asintóticas de los estimadores.¹⁸ Hasta el momento una buena parte de la literatura sobre estimadores de tamiz se refiere al caso de la regresión con datos transversales, aunque resultados como los de White (1990), y especialmente Chen y Shen (1998), permiten la extensión a ciertos tipos de series temporales.¹⁹ Por otra parte, la teoría permite la estimación simultánea de superficies de regresión y de sus derivadas así como otras características de interés, tales como elasticidades y otros funcionales econométricos.

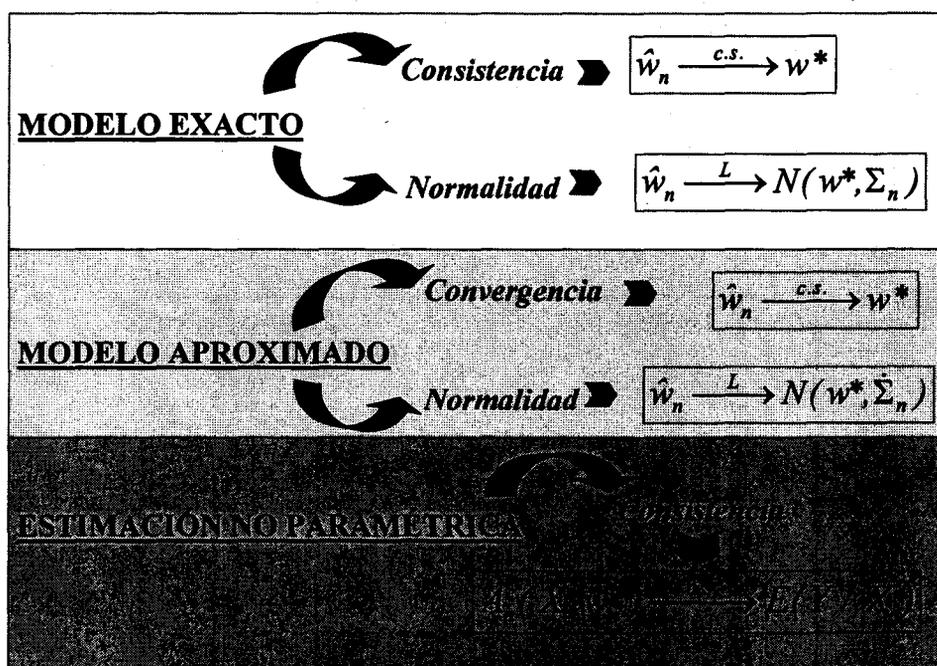


Figura 4: Aspectos inferenciales. Resumen de resultados generales (estimación).

En la práctica, la implementación del método de los tamices depende crucialmente del control de la velocidad de crecimiento de la complejidad de los modelos. Para ello se recurre a técnicas de validación cruzada o a criterios de información que penalizan el sobreajuste (Sin y White, 1996).

6. LIMITACIONES DE LOS MÉTODOS NO PARAMÉTRICOS

La principal ventaja que aportan los métodos paramétricos es la flexibilidad que hemos señalado, al no depender de una especificación del modelo. A cambio, hay que señalar algunas limitaciones importantes de los mismos:

¹⁸ Una interesante –y asequible– introducción a las ideas de la estimación no paramétrica, con especial énfasis en las redes neuronales, la constituye el trabajo de Geman *et al.* (1992).

¹⁹ Una observación importante es que la clase de modelos ARMA, que normalmente son vistos desde un punto de la estimación paramétrica, puede adaptarse para su uso en esquemas de estimación no paramétrica de tipo tamiz. El teorema de descomposición de Wold es un teorema de aproximación universal aplicable a los procesos estocásticos débilmente estacionarios, y el procedimiento de construcción de tamices se basaría en ir construyendo una sucesión de modelos ARMA de complejidad creciente (mayor número de retardos y/o términos MA) a medida que crece el tamaño de muestra.

1°. Su principal inconveniente viene dado por su menor eficiencia con relación a los métodos paramétricos. Es el fenómeno conocido como dilema sesgo-varianza o flexibilidad-eficiencia, que conlleva que los estimadores no paramétricos suelen requerir tamaños de muestra más elevados que los métodos clásicos o paramétricos. La razón intuitiva de este fenómeno es que los métodos no paramétricos deben extraer de los datos la información de contorno que en el caso de los métodos paramétricos viene aportada por el investigador durante la fase de especificación del modelo. Esta información relativa a especificación y la propia información de origen muestral son en gran medida sustitutivas, y *grosso modo* lo que no aporta el investigador debe extraerse de los datos, por lo que se requieren mayores tamaños muestrales. En esencia, los métodos no paramétricos cambian eficiencia por flexibilidad. Por este motivo, cuando se disponga de un buen modelo paramétrico, la recomendación general ha de ser siempre preferir éste antes que el uso de métodos no paramétricos. En problemas más complejos, y especialmente cuando se disponga de muestras de tamaño suficiente, puede ser ventajoso el uso de estos últimos.

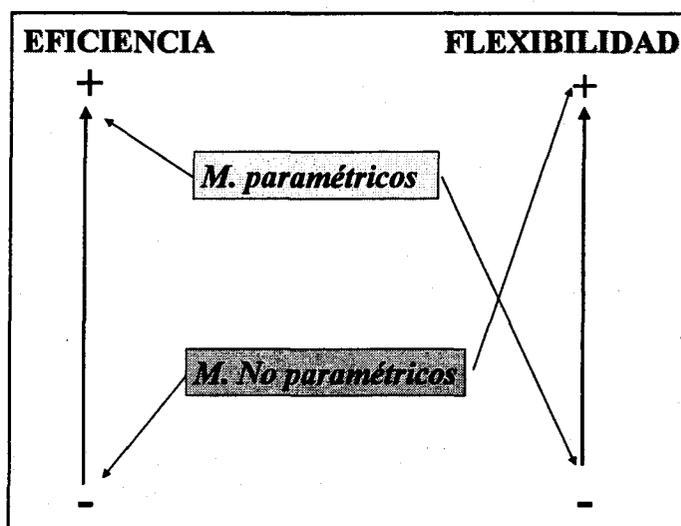


Figura 5: El dilema sesgo-varianza.

2°. Por otra parte, los métodos no paramétricos, a pesar de su flexibilidad, deben enfrentarse también a los demás problemas de los datos económicos: inexactitud de las observaciones (véase Morgenstern, 1970), escasez de datos, bajos ratios señal-ruido, presencia de cambios estructurales en los datos en serie temporal, observaciones anómalas y otros muchos. Además, los modelos fuertemente parametrizados, como es el caso de las técnicas flexibles, requieren extremar al máximo el control de la complejidad de los modelos.

3°. También se suele criticar el carácter de cajas negras computacionales de muchos métodos no paramétricos. Desde luego, los parámetros de una recta de regresión son más sencillos de interpretar que los de una RNA. En un nivel intermedio entre ambos se hallarían los estimadores de serie. Por otra parte, de poco sirve un modelo simple e interpretable si es sustancialmente sospechoso de ser incorrecto.

4°. La modelización con datos no estacionarios es un problema poco resuelto a nivel teórico. En el caso de los modelos no lineales se está buscando aún un concepto análogo

al de cointegración, aunque existen algunos resultados basados en el concepto de los atractores extraños, que serían la extensión natural de la cointegración al caso de los sistemas dinámicos no lineales. Por otra parte, hasta el momento presente los procesos de modelización no lineales, incluso en el caso no paramétrico, se basan en la inclusión en los modelos de variables que hayan sido previamente tratadas para conseguir estacionariedad y linealidad en varianza. Son poco conocidas las distorsiones provocadas por estos tratamientos en el caso general, aunque ha de presumirse que serán al menos del mismo orden que en el caso de la modelización ARIMA.

5º. Los métodos flexibles descritos están básicamente diseñados para tratar un tipo particular de no linealidades: a las no linealidades en media (deterministas), se le pueden añadir no linealidades en varianza, estocásticas. Los algoritmos recursivos para estimar redes recurrentes se pueden adaptar perfectamente a esta tarea, y sería interesante una exploración de las capacidades de modelos como las RNA recurrentes en el tratamiento conjunto de no linealidades en media y varianza.

6º Los métodos paramétricos ofrecen rutinas más estandarizadas para su uso y un software familiar para el usuario. Las técnicas flexibles son, al menos por ahora, menos cómodas.

7º. Finalmente, la distinción entre métodos paramétricos y no paramétricos no es siempre clara, como señalan Teräsvirta *et al.* (1994).

7. APLICACIONES A LA PREDICCIÓN DE LAS SERIES DE LA ECONOMÍA DE ASTURIAS

Las series económicas presentan características específicas que hacen en ocasiones muy difícil su adecuada modelización y su posterior utilización con fines predictivos. Se trata en general de series no estacionarias, con elevado nivel de ruido, frecuente presencia de cambios estructurales, observaciones anómalas y otros muchos fenómenos que hacen difícil la identificación de estructuras estables en las mismas. Esta misma variedad de comportamientos hace que en la práctica no exista una técnica de modelización y predicción universalmente óptima, que proporcione mejores resultados que las demás en todas las situaciones. Dado que la orientación básica de la predicción económica es eminentemente pragmática, el usuario debe disponer de una caja de herramientas equipada con una variedad suficiente de modelos.

A modo de ilustración de estas ideas, comentaremos brevemente los aspectos más relevantes de una serie de iniciativas que, dentro del equipo HISPALINK-Asturias, se vienen llevando a cabo con el objetivo de integrar las diversas metodologías flexibles en la predicción de las series de la economía asturiana.

7.1 Antecedentes y estado actual del tema

La predicción de las series económicas asturianas se viene realizando hasta la fecha mediante la metodología ARIMA. Se utiliza a tal efecto el programa PROYECTA desarrollado en Eviews que incorpora las distintas etapas del tratamiento Box-Jenkins de un buen número de series mensuales y trimestrales y la consiguiente evaluación de los resultados. La base de datos AsturDat de HISPALINK-Asturias contiene un elevado

número de series mensuales y trimestrales, muchas de las cuales son utilizadas como variables explicativas en diversos modelos econométricos regionales (Pérez *et al.*, 1994). Resulta de suma importancia cuidar la calidad de las proyecciones realizadas sobre dichas series, ya que condicionan las estimaciones y predicciones referidas a los principales agregados de la economía regional (principalmente PIB y empleo sectoriales)²⁰.

Entre los enfoques alternativos al problema de la predicción de las series de la economía asturiana, el planteamiento de modelos neuronales llevado a cabo recientemente (Landajo, 1999) para algunas series de precios y producción regionales permite apreciar algunas ventajas de esta metodología, que aconsejan un estudio de su comportamiento en una batería amplia de magnitudes. En principio, la variedad tipológica de las series económicas hace esperar que dentro de las series asturianas existan grupos más adecuadamente modelizables mediante metodologías ARIMA o VARMA y otros más adecuados para su tratamiento mediante modelos no lineales, tales como las redes neuronales o incluso los sistemas de reglas u otros procedimientos flexibles.

7.2 Metodología de trabajo para la modelización no lineal

En la construcción de modelos no lineales para las series económicas se requiere, aún más que en el caso lineal, de una estrategia o rutina de modelización, un procedimiento relativamente sistemático que sirva de guía al usuario. Para el caso de los modelos neuronales resulta de utilidad el esquema metodológico recogido en la tabla 1, propuesto en Landajo (2000). Se trata esencialmente de una extensión de la metodología Box-Jenkins al caso de los modelos neuronales²¹. Como hemos indicado más arriba, en lo que se refiere a su uso en la modelización de series temporales, las RNA pueden tratarse como una clase de modelos NARMAX.

La estimación de los modelos neuronales se puede llevar a cabo por medio de un software diseñado específicamente para el tratamiento de series económicas, que puede ejecutarse en el entorno Matlab, y que ha sido utilizado con resultados satisfactorios en aplicaciones previas.

7.2 Combinación de predicciones

Otro aspecto a considerar es la posibilidad de utilizar predictores combinados. Desde una óptica teórica, las ventajas de la combinación de predicciones aparecen recogidas en los trabajos de Bates y Granger (1969) y Wrinkler y Makridakis (1983). Bajo determinadas condiciones, dichos predictores combinados presentan *a priori* propiedades más satisfactorias que los predictores originales, por lo que su utilización abre una vía adicional de mejora. También han sido numerosos los esfuerzos empíricos dedicados por diversos autores al análisis de las predicciones combinadas.

²⁰El equipo HISPALINK-Asturias, bajo la dirección de Rigoberto Pérez Suárez, viene realizando de forma continuada tareas de modelización y predicción de series de la economía asturiana. Los principales resultados de las predicciones regionales aparecen recogidos en los informes semestrales "*Situación actual y perspectivas económicas de Asturias*" y "*Boletín de Coyuntura de Asturias*" (ambos disponibles en el web www.hispalink.org/asturias).

²¹ Es sabido que no existe rutina de modelización de las series temporales a salvo de limitaciones. En este sentido, nuestra propuesta comparte la mayoría de las ventajas y limitaciones de la metodología de Box y Jenkins (1970), y Box, Jenkins y Reinsel (1994).

Tabla 1: Fases del proceso de modelización de una serie temporal.

MODELIZACIÓN FLEXIBLE DE SERIES TEMPORALES
BÚSQUEDA DE TENDENCIA ESTOCÁSTICA/DETERMINÍSTICA Test de rachas. Contrastes de raíz unitaria.
TRANSFORMACIÓN A ESTACIONARIEDAD Log (Box-Cox). Diferenciación. Modelización o eliminación de tendencia determinista.
BÚSQUEDA DE TÉRMINOS AR Y NAR Términos a incluir (regulares/estacionales). Lineales/no lineales. Número de no linealidades (neuronas, reglas) (criterios de información, validación cruzada).
BÚSQUEDA DE ESTRUCTURAS RECURRENTES Lineales: Términos MA (en series diferenciadas). No lineales.
OBSERVACIONES ANÓMALAS Detección automática. Modelización (términos lineales con variables dicotómicas).
CAMBIOS ESTRUCTURALES Búsqueda, análisis. Modelización (ventanas móviles, muestras truncadas, variables explicativas).
DIAGNÓSTICOS DE RESIDUOS Estructuras no modelizadas, correlación serial (lineal/no lineal). Contrastes de normalidad.
DIAGNÓSTICOS DE PREDICCIÓN U de Theil, R^2 , E.A.M. Análisis de coherencia del modelo construido. Comparación con otros modelos (p.e. el paseo aleatorio con deriva). Posibilidad de combinación de predicciones.
USO PRÁCTICO

En el caso de las series regionales algunas aplicaciones de comparación de predicciones con métodos exponenciales, ARIMA y neuronales, así como de combinaciones de los mismos han sido llevadas a cabo recientemente (Moreno *et al.*, 2000). Aunque los análisis se circunscriben por el momento a un número limitado de series, los resultados parecen corroborar las ideas anteriores, en el sentido de que la incorporación de predicciones basadas en distintas técnicas (ARIMA, RNA, alisados) y su combinación ofrece en principio cierto margen de mejora frente a los enfoques de herramienta única.

7.3 Objetivos e implementación

La aplicación, a escala general, de las ideas anteriores incluye los siguientes aspectos:

Fase I: Análisis comparativo de las capacidades predictivas de las metodologías ARIMA y RNA en el ámbito de las principales series económicas asturianas. Como hemos señalado, esta cuestión no ha sido aún explorada salvo en algunas series aisladas

de ámbito nacional, y no existen análisis de esta clase para series regionales españolas, y menos aún análisis para grupos amplios de series, por lo que es de interés analizar baterías representativas.

Fase II: Establecimiento de una tipología de las principales series de la economía asturiana, en la que se clarifique qué series se predicen mejor con modelos lineales ARIMA y cuáles se adaptan mejor a esquemas no lineales, en particular los modelos neuronales.

Fase III: Análisis de las posibilidades que ofrece, para cada una de las series, la combinación de predicciones de ambas clases de modelos, como alternativa al uso de una sola categoría de modelos. Se incluye en este objetivo la elección de combinaciones óptimas de predictores.

Fase IV: Aplicación de los resultados obtenidos a la mejora de las predicciones de la economía asturiana (figura 6).

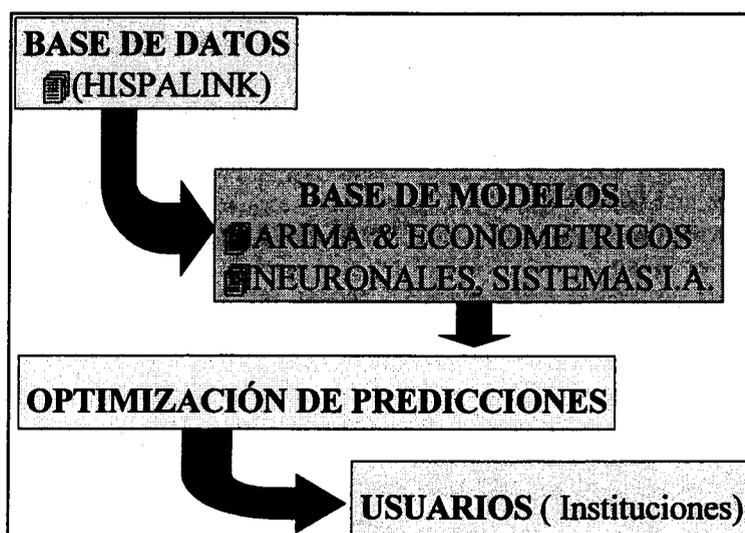


Figura 6: Proceso de generación de predicciones.

8. CONCLUSIONES Y OBSERVACIONES FINALES

En este trabajo se han presentado los aspectos fundamentales de las principales técnicas de regresión flexibles. Hemos señalado las importantes comunalidades existentes entre las diversas técnicas flexibles y el papel de la estimación de tamiz en el diseño y análisis de los procedimientos de estimación no paramétricos. La considerable amplitud y complejidad de la literatura técnica, unida a su diversidad y relativa novedad hacen que - con la notable excepción de la econometría financiera- los usuarios económicos de las técnicas estadísticas no hayan aprovechado aún de manera suficiente las ventajas de estos métodos.

Desde un punto de vista práctico, la relación óptima entre métodos paramétricos y no paramétricos ha de ser de complementariedad más que de competencia. La ausencia de técnicas universalmente dominantes aconseja al investigador o usuario disponer de una panoplia suficientemente amplia de herramientas de predicción. Una estrategia de trabajo útil puede ser comenzar siempre por utilizar los métodos más simples, que en

general serán los paramétricos, y recurrir a enfoques más sofisticados sólo en caso de necesidad. En este sentido, se han descrito algunos esfuerzos recientes de integración de métodos paramétricos y no paramétricos en el ámbito de la predicción regional.

REFERENCIAS BIBLIOGRÁFICAS

- Andrews, D. K. (1991): "Asymptotic normality of series estimators for nonparametric and semiparametric regression models", *Econometrica*, Vol. 59, Nº 2, March, pp. 307-345.
- Bates, J.M. y Granger, C.W.J. (1969): "The combination of Forecast", *Operational Research Quarterly*, Vol 20, n. 4, pp. 451-468.
- Box, G.E.P. ; Jenkins, G.M. (1970) *Time Series Analysis: Forecasting and control*. Holden-Day.
- Box, G.E.P., Jenkins, G.M. ; Reinsel, G.C. (1994) *Time Series Forecasting and control*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Chen, X.; Fan, Y. (1999): "Consistent hypothesis testing in semiparametric and nonparametric models for econometric time series", *Journal of Econometrics*, 91, pp. 373-401.
- Chen, X.; Shen, X. (1998): "Sieve Extremum Estimates for Weakly Dependent Data", *Econometrica*, Vol. 66, No. 2, March, pp. 289-314.
- Chen, X.; White, H. (1997): "Improved rates and asymptotic normality for nonparametric neural network estimators", *Discussion Paper 97-11, May*. University of California, San Diego.
- Cheng, B.; Titterton, D.M. (1994): "Neural Networks: A Review from a Statistical Perspective", *Statistical Science*, Vol. 9, nº 1, pp. 2-54.
- Chui, Ch. K. (1988): *Multivariate Splines*. SIAM, Philadelphia, Pennsylvania.
- Connor, J.T.; Martin, R.D.; Atlas, L.E. (1994): "Recurrent Neural Networks and Robust Time Series Prediction", *IEEE Transactions on Neural Networks*, Vol. 5, No. 2, March, p.p. 240-254.
- Cottrell, M; Girard, B.; Girard, Y.; Mangeas, M.; Muller, C. (1995): "Neural Modeling for Time Series: A Statistical Stepwise Method for Weight Elimination", *IEEE Transactions on Neural Networks*, Vol. 6, Nº 6, November, pp. 1355-1364.
- Davidson, R.; Labys, W.C.; Lesourd, J.B. (1998): "Wavelet Analysis of Commodity Price Behaviour", *Computational Economics*, Vol. 11, Nos. 1-2, April, pp. 103-128.
- Day, R.H. (1994): *Complex Economic Dynamics* (Vol. 1), The MIT Press, Cambridge, MA.
- DeVore, R.A.; Lorentz, G.G. (1993): *Constructive Approximation*. Springer-Verlag, Berlin.
- Domowitz, I.; White, H. (1982): "Misspecified models with dependent observations", *Journal of Econometrics*, 20, pp. 35-58.
- Elbadawi, I.; Gallant, A. R.; Souza, G. (1983): "An elasticity can be estimated consistently without a priori knowledge of functional form", *Econometrica*, Vol. 51, Nº 6, November, pp. 1731-1751.
- Fix, E.; Hodges, J.L. (1951): "Discriminatory analysis nonparametric estimation: consistency properties", *Report No. 4, Project No. 21-49-004*, USAF School of Aviation Medicine, Randolph Field, Texas.
- Friedman, J.H.; Stuetzle, W. (1981): "Projection Pursuit Regression", *Journal of The American Statistical Association* 76, pp. 817-823.
- Gallant, A.R. (1987): "Identification and consistency in semi-nonparametric regression", en Truman F. Bewley (Ed.), *Advances in econometrics fifth world congress, Vol. I* (pp. 145-170). New York: Cambridge University Press.
- Gallant, A.R.; White H. (1992): "On Learning of the Derivatives of an Unknown Mapping With Multilayer Feedforward Networks", *Neural networks*, Vol. 5, pp. 129-138.
- Geman, S.; Bienenstock, E.; Doursat, R. (1992): "Neural Networks and the Bias/Variance Dilemma", *Neural Computation*, 4, pp. 1-58.
- Geman, S.; Hwang, C.R. (1982): "Nonparametric Maximum Likelihood Estimation by The Method of Sieves", *The Annals of Statistics*, Vol. 10, nº 2, pp. 401-414.
- Granger, C.W.J. (1990): "General Introduction", en *Modelling Economic Time Series* (C.W.J. Granger Ed.), Oxford University Press, Oxford.
- Grenander, U. (1981): *Abstract Inference*. Wiley, New York.
- Grossberg, S. (1988): "Nonlinear Neural Networks: Principles, Mechanisms, and Architectures", *Neural Networks*, Vol. 1, pp. 17-61.
- Haavelmo, T. (1944): "The Probability Approach in Econometrics", *Econometrica*, Vol. 12, Supplement.
- Harvey, A.C. (1989): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- Hong, B.Y. ; White, H. (1995): "Consistent specification testing via nonparametric series regression", *Econometrica*, Vol. 63, Nº 5, September, pp. 1133-1159.

- Huber, P.J. (1964): "Robust Estimation of a Location Parameter", *Annals of Mathematical Statistics*, 35, pp. 73-101.
- Huber, P.J. (1967): "The Behaviour of Maximum Likelihood Estimates Under Nonstandard conditions", *Proceedings of The Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, 1, pp. 221-233.
- Johnson, J.D.; Whinston, A.B. (1994): *Advances in artificial intelligence in economics, finance, and management*. Volume 1. JAI PRESS INC., Greenwich, Connecticut, 1994.
- Kendall, M.; Ord, K. (1990): *Time series*. Edward Arnold. Sevenoaks, Kent.
- Kosko, B. (1992): *Neural networks and fuzzy systems. A dynamical systems approach to machine intelligence*. Prentice Hall International, Englewood Cliffs, N.J.
- Kuan, C.M., Hornik, K., White, H. (1990): "Some Convergence Results for Learning in Recurrent Neural Networks", UCSD Department of Economics Discussion Paper.
- Kuan, C.M.; White, H. (1994a): "Artificial Neural Networks: An Econometric Approach", *Econometric Reviews*, 13 (1), pp. 1-91.
- Kuan, C.M.; White, H. (1994b): "Adaptive Learning with Nonlinear Dynamics Driven by Dependent Processes", *Econometrica*, Vol. 62, No. 5, September, pp. 1087-1114.
- Landajo, M. (1999): *Modelos neuroborrosos para la predicción económica*. Tesis Doctoral, Universidad de Oviedo, Departamento de Economía Aplicada.
- Landajo, M. (2000): "Neural and Fuzzy Models for Economic Forecasting. An Econometric View and Some Practical Experience", *Fuzzy Economic Review*, Vol. 5, no. 1, May, pp. 3-28.
- Lapedes, A.; Farber, R. (1987): "Nonlinear signal processing using neural networks", *Technical Report LA-UR-87-2662*, Los Alamos National Laboratory, Los Alamos, N.M.
- López, A.J.; Moreno, B. (1999): "Evaluación de predicciones a partir de medidas de información. Nuevas Alternativas", Comunicación, *XIII Reunión Asepelt-España*, Actas De La XIII Reunión Asepelt-España, Burgos.
- López, A.J.; Pérez, R. (1998): "Análisis de la coyuntura regional. Técnicas de estimación y predicción", *Revista Asturiana de Economía*, Volumen: 11, Páginas 71-91.
- Maasoumi, E. (1994): "Artificial Neural Networks for Some Macroeconomic Series: A First Report", *Econometric Reviews*, Vol. 13, Nº 1, pp. 105-122.
- Marcet, A.; Sargent, T.J. (1989): "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, 48, pp. 337-368.
- Moreno, B., López, A.J.; Landajo, M. (2000): "Técnicas de predicción en series temporales y medidas de evaluación", *Actas de XIV Jornadas Asepelt-España*, Oviedo.
- Morgenstern, O. (1970): *Sobre la exactitud de las observaciones económicas*. Ed. Tecnos, S.A., Madrid.
- Pan, Z.; Wang, X. (1998): "A Stochastic Nonlinear Regression Estimator Using Wavelets", *Computational Economics*, Vol. 11, Nos. 1-2, April, pp. 89-102.
- Parzen, E. (1962): "On the estimation of probability density and mode", *Annals of Mathematical Statistics* 33, pp. 1065-1076.
- Parzen, E. (1979): "Nonparametric statistical data modelling", *Journal of The American Statistical Association* 74, pp. 105-131.
- Pérez, R.; López, A.J.; Caso, C.; Río, M.J.; Hernández, M. (1994): "Mecastur. Modelo econométrico para la economía asturiana", *Cuadernos Aragoneses de Economía*, Volumen: 4, pp. 273-292.
- Poggio, T.; Girosi, F. (1989): "A Theory of Networks for Approximation and Learning", Artificial Intelligence Laboratory, MIT, AI Memo nº 1140, July.
- Pötscher, B.M.; Prucha, I.R. (1991): "Basic Structure of The Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts", *Econometric Reviews*, Vol. 10, Nº 2, pp. 125-216.
- Pötscher, B.M.; Prucha, I.R. (1991): "Basic Structure of The Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part II: Asymptotic Normality", *Econometric Reviews*, Vol. 10, Nº 3, pp. 253-325.
- Rosenblatt, M. (1956): "Remarks on some nonparametric estimates of density function", *Annals of Mathematical Statistics* 27, pp. 832-837.
- Shen, X. (1997): "On Methods of Sieves and Penalization" *The Annals of Statistics*, Vol. 25, Nº 6, pp. 2555-2591.
- Sin, C.Y.; White, H. (1996): "Information criteria for selecting possibly misspecified parametric models", *Journal of Econometrics*, 71, pp. 207-225.
- Swanson, N.R.; White, H. (1995a): "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks", *Journal of Business and Economic Statistics*, July, Vol. 13, nº 3, pp. 265-275.

- Swanson, N.R.; White, H. (1995b): "A Model-Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks", UCSD, Research Group for Econometric Analysis, *Discussion Paper*.
- Teräsvirta, T. (1994): "Specification, estimation and evaluation of smooth transition autoregressive models", *Journal of The American Statistical Association*, Vol. 89, pp. 208-218.
- Teräsvirta, T.; Anderson, H.M. (1992): "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models", *Journal of Applied Econometrics*, 7, S119-S136.
- Teräsvirta, T.; Tjøstheim, D.; Granger, C.W. (1994): "Aspects of modelling nonlinear time series", en (R.F. Engle & D.L. McFadden, eds.) *Handbook of Econometrics*, vol. IV, cap. 48, pp. 2917-2957. Elsevier Science, New York.
- Tong, H. (1990): *Non-linear time series. A dynamical system approach*. Oxford University Press, Oxford.
- Ullah, A. (1988): "Non-parametric estimation of econometric functionals", *Canadian Journal of Economics*, Vol. XXI, N° 3, August, pp. 625-658.
- White, H. (1982): "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, Vol. 50, N° 1, pp. 1-25.
- White, H. (1981): "Consequences and Detection of Misspecified Nonlinear Regression Models", *Journal of The American Statistical Association*, Vol. 76, N° 374, June, pp. 419-433.
- White, H. (1987): "Some Asymptotic Results for Back-propagation", *Proceedings of The First International Conference on Neural Networks*, San Diego. New York: IEEE Press, pp. III:261-266.
- White, H. (1989a): "An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks", *Proceedings of The International Joint Conference on Neural Networks*, Washington, DC. Vol. 2. IEEE Press, New York, pp. 451-455.
- White, H. (1989b): "Learning in Artificial Neural Networks: A Statistical Perspective", *Neural Computation*, vol. 1, n°4, Winter, pp. 425-464.
- White, H. (1989c): "Neural Network Learning and Statistics", *AI Expert*, December, pp. 48-52.
- White, H. (1989d): "Some Asymptotic Results for Learning in Single Hidden Layer Feedforward Artificial Neural Networks", *Journal of The American Statistical Association*, December, Vol. 84, n° 408, (Theory and Methods), pp. 126-131.
- White, H. (1990): "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings", *Neural Networks*, Vol. 3, pp. 535-549.
- White, H. (1994): *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York.
- Wooldridge, J.M. (1994): "Estimation and inference for dependent processes", en (R.F. Engle & D.L. McFadden, eds.) *Handbook of Econometrics*, vol. IV, cap. 45, Elsevier Science B.V., pp. 2639-2738.
- Wrinkler, R.L. Makridakis, S. (1983): "The combination of Forecast", *Journal of The Royal Statistical Society*, Serie A, 146, Part.2, pp. 150-157.
- Xu, L.; Krzyzak, A.; Yuille, A. (1994): "On Radial Basis Function Nets and Kernel Regression: Statistical Consistency, Convergence Rates, and Receptive Field Size", *Neural Networks*, Vol. 7, n° 4, pp. 609-628.
- Zadeh, L. (1973): "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-3, n° 1, January 1973, pp. 28-44.