



ANÁLISIS Y PREDICCIÓN DE LA SERIE DE TIEMPO DEL PRECIO EXTERNO DEL CAFÉ COLOMBIANO UTILIZANDO REDES NEURONALES ARTIFICIALES

Ismael García Martín

Facultad de Ciencias, Pontificia Universidad Javeriana.
e-mail: ismael.garcia@javeriana.edu.co

RESUMEN

En el presente artículo se busca diseñar un modelo no lineal para el análisis y predicción de la serie de tiempo del precio externo del café colombiano utilizando *Redes Neuronales Artificiales*. El objetivo es comparar los alcances y limitaciones de este modelo frente a un modelo clásico de predicción lineal ARIMA.

Palabras clave: Serie de tiempo, ARIMA, Red Neuronal Artificial.

ABSTRACT

This article intends to design a non linear model for analysis and prediction of the time-series concerning the external price of Colombian coffee using Artificial Neural Networks. The main purpose is to compare the realm and limitations of proposed model and the linear prediction classic model ARIMA.

Key words: Time-series, ARIMA, Artificial Neural Networks.

INTRODUCCIÓN

El precio internacional del café presenta variaciones importantes año a año; los factores a los que se ha asociado la inestabilidad del precio externo se refieren a condiciones climáticas, acuerdos internacionales y acumulación de existencias por parte de los países consumidores y productores, los cuales a su vez generan reducciones de oferta. Todo esto genera un riesgo importante para aquellos países en desarrollo, como Colombia, cuyo sector exportador se encuentra concentrado, en parte, en este tipo de bien. La posibilidad de predecir el comportamiento del precio internacional

del café, se convierte, pues, en una herramienta fundamental en el manejo de la política cafetera.

Modelos matemáticos han sido diseñados para el tratamiento de este tipo de problemas. El más conocido es el Proceso Autorregresivo Integrado de Media Móvil (ARIMA), comúnmente conocido como la metodología de Box-Jenkins (1976). Este método de predicción, que se utiliza muy frecuentemente, se fundamenta en el supuesto implícito de la linealidad del sistema que generan la trayectoria de las variables.

REDES NEURONALES ARTIFICIALES

Una red neuronal es un modelo computacional, que recientemente ha generado un gran entusiasmo, tanto en Colombia como en el resto del mundo, por sus capacidades potenciales, la cual posee una arquitectura masivamente paralela que trata de emular la estructura neuronal del cerebro humano, consta de varias unidades sencillas que trabajan en paralelo sin un control central.

Existe una gran variedad de tipos de estructuras de redes neuronales; la principal diferencia está entre las redes de *prealimentación* y las *redes recurrentes*. En las primeras las conexiones son unidireccionales y no hay ciclos. En las recurrentes, las conexiones pueden formar topologías arbitrarias. Las redes de prealimentación con niveles fueron estudiadas por primera vez, por Rosenblatt y otros más, a finales de la década de los cincuenta bajo el nombre de *perceptrones*. Si bien se sometieron a estudio redes de todos los tamaños y topologías, el único elemento de aprendizaje efectivo en esa época fueron las redes de un solo nivel, por lo que en ellas fue concentrada la mayor parte de la atención. En la figura 1 mostramos un ejemplo muy sencillo de una red de prealimentación con dos niveles.

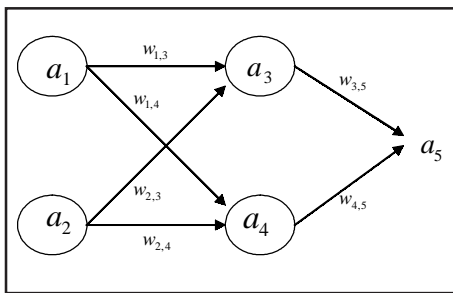


FIGURA 1. Red de prealimentación de dos niveles; con dos entradas, dos nodos ocultos y un nodo de salida.

Actualmente el nombre perceptrón es empleado como sinónimo de una red de prealimentación con un solo nivel y es una función de la forma:

$$y_t = A(W_t \cdot a_t), \quad (2.1)$$

donde W_t y a_t son vectores cada uno de tamaño $k \times 1$ y A es la función de activación. Dada una muestra de entrenamiento $\{y_t, a_t\}_{t=1}^r$, donde t es un individuo particular, el problema de entrenamiento del perceptrón es elegir W que minimice $\sum_{t=1}^r (y_t - A(W_t \cdot a_t))^2$. Esto es un problema de Mínimos Cuadrados No Lineales.

Una red neuronal de prealimentación con una capa oculta se puede describir por las siguientes dos ecuaciones:

$$y_t = \theta_0 + \sum_{j=1}^q \theta_j b_{jt}, \quad b_{jt} = A\left(\sum_{i=1}^k w_{ji} a_{it}\right). \quad (2.2)$$

La segunda ecuación describe la salida b_{jt} de la unidad oculta j , que es simplemente un perceptrón. La primera ecuación genera las salidas y_t de la red aplicando una función que pondera los vectores a_t , de tamaño $k \times 1$, de las salidas de las unidades ocultas. Los parámetros de la red son los pesos w_{ji} y θ_j . La red se puede representar en forma compacta como:

$$y_t = \theta_0 + \sum_{j=1}^q \theta_j A(W_j \cdot a_t) \quad (2.3)$$

En la literatura se encuentran dos direcciones acerca de esta clase de modelos: *representación* y *estimación*. Sobre representación o aproximación se limitan la clase de funciones $g: X \rightarrow Y$ que pueden ser arbitrariamente bien aproximadas por el modelo (2.3). Hornik, Stinchcombe y White (1989) demuestran que una red neuronal de tres capas tiene la capacidad de aproximar vir-



tualmente cualquier función de interés con una cantidad “suficiente” de unidades en la capa oculta, las cuales están provistas de funciones de activación de tipo sigmoïdal. El parámetro que controla la precisión de la aproximación es q , el número de unidades ocultas. Para un valor dado de q , la mejor aproximación a una función $g(x)$ es determinada por los valores de (θ, w) que minimizan la norma:

$$\left\| g(x) - \theta_0 + \sum_{j=1}^q \theta_j A(W_j \cdot a) \right\|^2, \quad (2.4)$$

donde $\|\cdot\|^2$ es la norma L_2^1 . La literatura sobre aproximación supone que se puede seleccionar q suficientemente grande para poder encontrar (θ, w) que hace que esta norma sea tan pequeña como se quiera.

El problema de *estimación* ocurre cuando tenemos dada una muestra, y un modelo particular de la forma (2.3), con q fijo, y deseamos estimar los parámetros, que en este caso son los θ 's y los w 's. Esto es una versión de un problema de regresión no lineal. El modelo para ser estimado se rescribe como

$$y_t = g(a_t, \beta) + \varepsilon_t \quad (2.5)$$

donde hemos reunido los parámetros θ y w en el vector β . Podemos proceder a la estimación utilizando una variedad de algoritmos basados en los métodos de aproximación estocástica.

Predicción del precio externo del café colombiano

El análisis clásico de regresión basado en información de series de tiempo supone implícitamente que la serie de tiempo en la cual se basa es estacionaria² o que puede convertirse en estacionaria mediante alguna transformación apropiada. De no ser así, el procedimiento convencional de prueba de hipótesis, basado en las pruebas t , F , ji -

cuadrado y otras similares, será de dudosa aceptación.

Es nuestro interés trabajar aquí con la serie de tiempo del precio promedio mensual del café colombiano en la Bolsa de Valores de Nueva York, desde enero de 1913 hasta julio de 2001, que en adelante llamaremos PC; el horizonte de predicción se limitó a los últimos siete meses: enero a julio de 2001. La gráfica de la serie PC se presenta en la figura 2. Un método muy popular de diseñar modelos para el análisis y la predicción de series de tiempo estacionarias es el ARIMA (metodología de Box-Jenkins). Como la serie PC es no estacionaria, se toma la serie de la variación porcentual de PC (que llamaremos VPC), que sí es estacionaria. La gráfica de la serie VPC se muestra en la figura 3. Comparada con la serie original PC dada en la figura 2 no exhibe tendencia alguna.

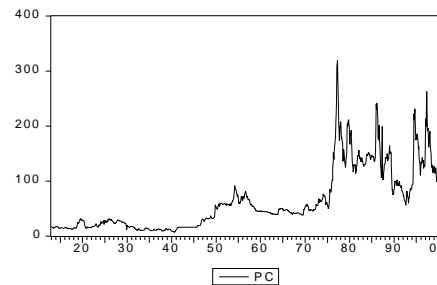


FIGURA 2. PC: precio externo del café colombiano, desde enero de 1913 hasta julio de 2001 (US\$ centavos por libra de 453.6 gramos).

¹ Si f es una función de valor real para la cual la integral de Lebesgue $\int |f|^p d\mu$ es finita, se define la norma L_p , $p \geq 1$, de la manera siguiente:

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}$$

² Si una serie de tiempo es estacionaria, su media, su varianza y su autocovarianza (en los diferentes rezagos) permanecen iguales sin importar el momento en el cual se midan.

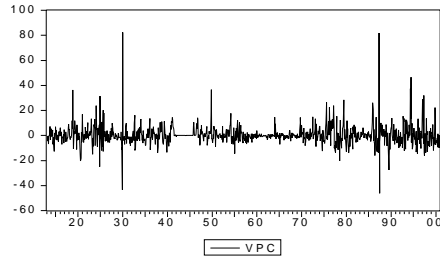


FIGURA 3. Gráfica de la serie VPC

Un modelo ARIMA, para la serie estacionaria VPC, estimado es:

$$VPC_t = 0.47 + 0.054VPC_{t-1} - 0.108VPC_{t-15} - 0.073VPC_{t-26} + 0.082VPC_{t-34} \quad (3.1)$$

$(0.25) (0.031) \quad (0.031) \quad (0.031) \quad (0.031)$
 $DW=2.009 \quad F=6.9 \quad S_{ARIMA}=7.843999$

donde los valores entre paréntesis son los errores estándar de los parámetros estimados y S_{ARIMA} es el error estándar del modelo (3.1). Los residuos del modelo son ruido blanco.

Ahora se presentan las Redes Neuronales Artificiales (RNA) como un método, no lineal, para el análisis y predicción de la misma serie VPC; el modelo especificado se debe estimar por NLLS. Bajo ciertas condiciones tales como estacionariedad de las variables, los estimadores de los parámetros son consistentes y asintóticamente normales.

Se empleó una red de prealimentación, de dos niveles con cuatro entradas, 11 unidades ocultas y una unidad de salida. Para la capa de entrada se tomaron los cuatro rezagos estadísticamente significativos al 95% con base en los coeficientes de autocorrelación parcial; es decir, las entradas son las variaciones porcentuales del precio del café en los rezagos t-1, t-15, t-26 y t-34; se eliminaron tres unidades o nodos de la capa intermedia y sus respectivas conexiones, a igual que otros enlaces entre

neuronas. Para todos los nodos de la red se utilizó como función de activación, la función sigmoide ($s(x) = \frac{1}{1+e^{-x}}$). La red resultante se muestra en la figura 4. La estimación de los parámetros (pesos) de la RNA, se realizó por el método de mínimos cuadrados no lineales utilizando el paquete econométrico EViews versión 3.0. La regresión no lineal estimada es de la forma, (3.2):

$$VPC_t = C(1) + \sum_{j=1}^{11} C(2+6(j-1)) * LOGIT(C(6(j-1) + i + 2) \cdot VPC_{t-d} + C(1+6j)) + e_t,$$

donde, $LOGIT(x) = s(x)$; $C(k)$, $k=1, \dots, 67$ son los θ 's y los w 's, los parámetros o pesos, del modelo (3.2), $i=1, 2, 3, 4$ y $d=1, 15, 26, 34$. El resultado de la estimación y el diagnóstico del modelo no se muestran (pero están disponibles para verificación). El error estándar del modelo (3.2) es $S_{RNA} = 6.930715$ y los residuos estimados de (3.2) son puramente aleatorios.

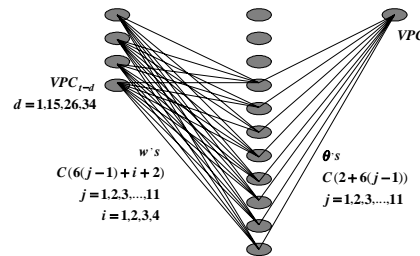


FIGURA 4. Topología de la RNA para la serie de VPC.

Para obtener la predicción de PC, se deshace la transformación que se utilizó para obtener las variaciones porcentuales; ésta se muestra en la figura 5. El promedio del precio real observado del café en dicho período es US\$75.72 centavos por libra y el promedio del precio pronosticado, en este mismo período, es US\$66.28 centavos

por libra para el modelo ARIMA y US\$70.60 centavos por libra para el modelo de RNA; el error promedio del pronóstico fue una subestimación de US\$9.44 y US\$5.12 centavos por libra, respectivamente, durante estos siete meses. En la figura 6 se muestra el comportamiento de las predicciones del modelo ARIMA y las del modelo de RNA, frente a la serie de los valores reales observados de VPC, durante el período de enero a julio de 2001.

Período	PC	PCF (ARIMA)	PCF (RNA)
2001:01	75.33	67.76467205	69.04304664
2001:02	76.7	66.02868088	69.47946711
2001:03	76.94	64.96218619	69.54728272
2001:04	78.25	65.71688594	69.6362823
2001:05	81.31	65.92901633	71.12129607
2001:06	72.24	66.86484416	71.94418531
2001:07	69.28	66.71766936	73.41427076

FIGURA 5. Valores reales observados de PC y los valores de su predicción PCF, entre enero y julio de 2001, con los modelos ARIMA y el de RNA.

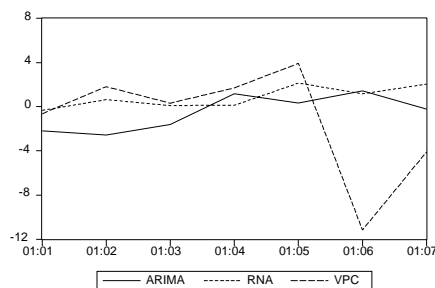


FIGURA 6. Predicciones para la serie de la variación porcentual del precio del café, VPC.

La predicción generada por la red neuronal fue mejor que aquella derivada del modelo ARIMA. Los resultados de la RNA estuvieron, en promedio, más cercanos en relación con la predicción del método clásico, aun-

que el último valor se encuentra lejano. No obstante, se observa que la RNA interpretó en forma correcta la tendencia de la serie VPC.

Un criterio para comparar estos dos modelos, se basa en la relación entre la varianza de los errores del modelo no lineal y la de los errores del modelo lineal. En general, se considera que dicha relación debe ser menor que 0.9 para que exista una diferencia entre el empleo de un modelo no lineal y uno lineal.

La relación entre la varianza del error del modelo de RNA (no lineal) y la del error del modelo ARIMA (lineal) es 0.78069 ($(S_{RNA})^2 / (S_{ARIMA})^2$). Significa esto que el modelo de RNA reduce la varianza del error del modelo ARIMA en, aproximadamente, un 22%. De acuerdo con el criterio, el modelo de RNA es superior frente al modelo ARIMA; es decir, el precio externo del café colombiano, se puede analizar y predecir con un modelo econométrico no lineal.

CONCLUSIONES

Gracias a las características de flexibilidad y adaptabilidad las Redes Neuronales Artificiales nos presentan una alternativa, no lineal, para analizar y predecir la serie de tiempo del precio externo del café colombiano. Sin embargo, como lo han demostrado algunos resultados, el uso de las redes neuronales no necesariamente desplazará los métodos clásicos de predicción, sino que al contrario, éstas se pueden convertir en procedimientos complementarios para un análisis más riguroso de las series de tiempo económicas.

En este trabajo se utilizaron Redes Neuronales Artificiales de prealimentación con una capa oculta. Como el algoritmo de propagación inversa es un algoritmo, diseñado para minimizar el error cuadrático medio entre la salida de la red y la salida de-



seada, requiere de funciones no lineales continuas y diferenciables; por ello se utilizó la función sigmoide como función de activación de la red ya que ésta satisface dichas condiciones. Sin embargo, la utilización de otros tipos de redes y su mayor profundización permitirán un mayor entendimiento de la aplicabilidad de las Redes Neuronales Artificiales al análisis y predicción de las series de tiempo económicas.

La ventaja de aplicar esta técnica radica en que, a diferencia de los modelos econométricos tradicionales, los cuales intentan “ajustar los datos a un modelo”; las Redes Neuronales Artificiales “fabrican un modelo que se ajusta a los datos”.

Con el fin de tener en cuenta otras variables importantes en la predicción del precio externo del café colombiano en la bolsa de valores de Nueva York, tales como, condiciones climáticas, acuerdos internacionales, acumulación de existencias, tasa de devaluación, etc., se recomienda la inclusión de estas variables en una Red Neuronal Artificial (RNA), ya que con base en lo aquí encontrado, ello augura resultados muy promisorios. Además podría compararse la RNA con un modelo lineal de vectores autorregresivo (VAR) encontrando, posiblemente, resultados a favor de aquella.

En el campo específico de las ciencias económicas, la posibilidad de predecir el comportamiento del mercado se convierte en una herramienta fundamental en el análisis econométrico y en el proceso de toma de decisiones. La información recolectada a través de las series de tiempo y los métodos de predicción pueden ser usados para prevenir posibles comportamientos no deseados en el mercado y tomar medidas que favorezcan la estabilidad del mismo.

Una de las principales características del precio externo del café es su volatilidad. Este comportamiento puede ser un buen

indicador de que los precios pueden no ser el resultado de un modelo lineal sino por el contrario, las asimetrías en las fluctuaciones y su gran volatilidad son un indicio de la presencia de no linealidades.

Existen problemas que no se han podido solucionar en este momento por medio de modelos matemáticos y que requieren de métodos no tradicionales de cálculo para obtener una respuesta. La gran mayoría de las situaciones económicas tienen un comportamiento no lineal, lo que sugiere utilizar métodos que los trabajen tal como son o que por medio de una transformación emulen un comportamiento lineal. Las redes neuronales al ser, por definición, no lineales y al comportarse como un sistema dinámico se presentan como una solución a este tipo de problemas.

LITERATURA CITADA

- CHO, IN-KOO Y SARGENT, THOMAS J. *Neural networks for encoding and adapting in dynamic economies. Handbook of Computational Economics*, vol. 1, Elsevier Science, 1996.
- GUJARATI, DAMODAR N. *Econometría*, Mc Graw-Hill, Bogotá, 1997, 717-735.
- HAMANN, FRANZ A. ¿Puede explicarse el precio externo del café con un modelo econométrico no lineal? *Desarrollo y Sociedad* No. 38. Bogotá: Universidad de los Andes, 1996.
- RUSSELL, STUART. *Inteligencia artificial, un enfoque moderno*. Prentice Hall, University of California, 1996.
- TORRES, LUZ GLORIA; HERNÁNDEZ, GERMÁN Y NIÑO, LUIS. *Redes neuronales*. 10° Coloquio Distrital de Matemáticas y Estadística. Universidad Nacional de Colombia, 1993.
- Revista del Banco de la República*. Principales Indicadores Económicos 1923-1997. Banco de la República (Colombia). 314-317.

Recibido: 11-05-2003

Aceptado: 08-08-2003