

## Capítulo 2

### **Series de tiempo financieras y pronóstico de rendimientos futuros.**

Una serie de tiempo financiera es una secuencia de observaciones (vectores característica) con las características de algunos activos financieros (como el precio  $P_t$ , retorno  $R_t$ , ganancia  $G_t$ , etc.) definidas en el tiempo discreto donde el índice de tiempo ( $t=1, \dots, N$ ) corresponde a una hora o un día de negociación, o una semana, un mes, etc. El precio de un activo,  $P_t$ , es una característica esencial dado dos precios diarios consecutivos,  $P_{t-1}$  y  $P_t$ . Otra característica financiera es el retorno definiéndose de la siguiente manera:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (1)$$

Donde  $-1 \leq R_t$

En nuestros experimentos, el precio de retorno acumulado es utilizado como una característica temporal por el hecho de que se requiere un rango dinámico más pequeño que hace que sea más fácil de cuantificar. Un rendimiento positivo,  $R_t > 0$  representa un aumento en el precio de los activos, mientras que un rendimiento negativo,  $R_t < 0$  significa una caída de precios. Otra característica de los activos en relación con el retorno es la de un periodo de aumento de  $G_t$  que viene dada por:

$$G_t = \frac{P_t}{P_{t-1}} = 1 + R_t \quad (2)$$

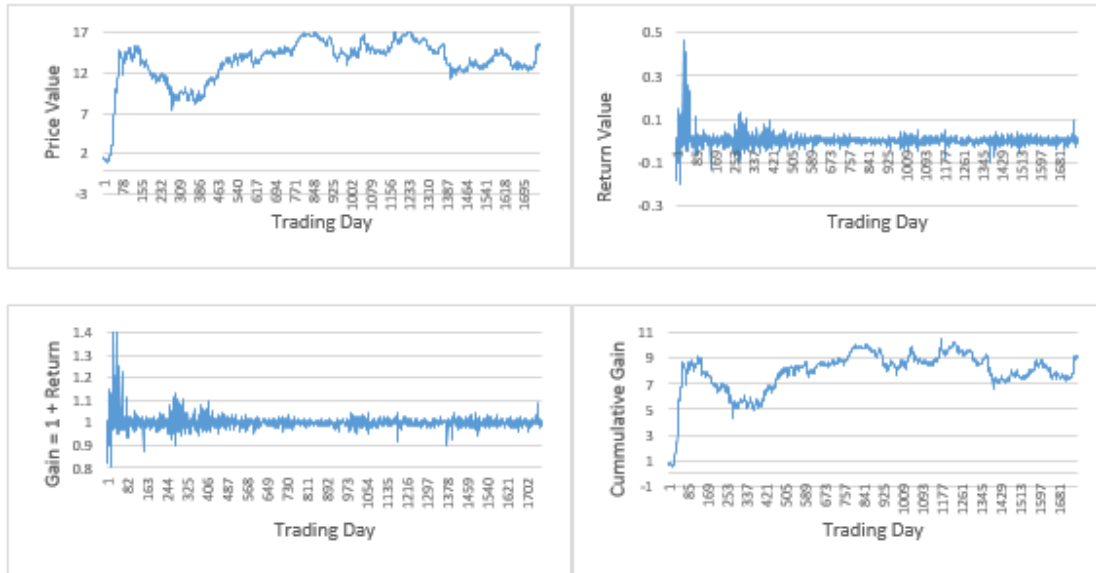
Un retorno o precio de incremento positivo implica una ganancia mayor que uno, mientras que un rendimiento negativo o caída de los precios presenta una ganancia menor que uno. La predicción de series de tiempo financieras consiste en estimar un futuro precio de retorno,  $R_{t+n}$ , con  $n$  días comerciales de anticipación mediante el procesamiento de un conjunto de las últimas observaciones de retorno  $R_t, R_{t-1}, \dots, R_{t-\ell+1}$ .

El  $P_c$  precio de retorno acumulado después de un periodo de días de negociación  $n$ , viene dado por el producto del precio inicial de los  $P$  activos y la ganancia acumulada de acuerdo con:

$$P_c = P (1 + R_1) (1 + R_2) \dots (1 + R_n) \quad (3)$$

El retorno no sólo se utiliza como característica sino que también sirve para medir el rendimiento de un predictor financiero artificial. Un predictor financiero es eficiente si proporciona un inversor con ganancia acumulada superior a la obtenida por no usarlo. La consecuencia de esta predicción es que el inversor decide vender sus acciones a los días en que se espera una caída de los precios, por lo que la rentabilidad obtenida en esos días es cero  $R_t = 0$  en lugar de negativo con una ganancia de  $1 + R_t = 1$ . Por lo tanto, el resultado acumulado en el periodo de cuatro días es  $(1 + R_2)(1 + R_4)$ , que se espera sea una predicción exacta y más alto que el índice original  $(1 + R_1)(1 + R_2)(1 + R_4)(1 + R_4)$ . Una curva de ganancia de retorno muestra la evolución de la ganancia de rendimiento a medida que aumenta el tiempo y el objetivo final de un predictor financiero artificial es generar una curva de ganancia acumulada superior a la curva de índice. Cuanto mayor es la precisión de la predicción mayor será la rentabilidad obtenida. La figura 1 muestra un conjunto de series de tiempo financieras

que incluye curvas de precio de las acciones, el valor de retorno, la ganancia en cada día de entrenamiento y la ganancia acumulada.



*Figura 1 Series de tiempo financieras*

La predicción de series de tiempo financieras es un proceso de dos etapas:

- La formación del modelo para aprender los parámetros que maximizan la probabilidad de ocurrencia de una secuencia de observación de los valores de retorno.
- La predicción del valor de retorno para el siguiente día de negociación que viene.

Después de la predicción las dos etapas se repiten alternativamente mediante el uso de una secuencia ajustada de las observaciones que suma el valor de retorno observado más nuevo y deja caer el valor de retorno observado más antiguo de la secuencia.

La longitud de las secuencias de observaciones puede tomar valores diferentes tales como 5, 20 o 50 días de negociación. La figura 2 muestra las dos etapas de un modelo predictivo aplicado a series de tiempo financieras.

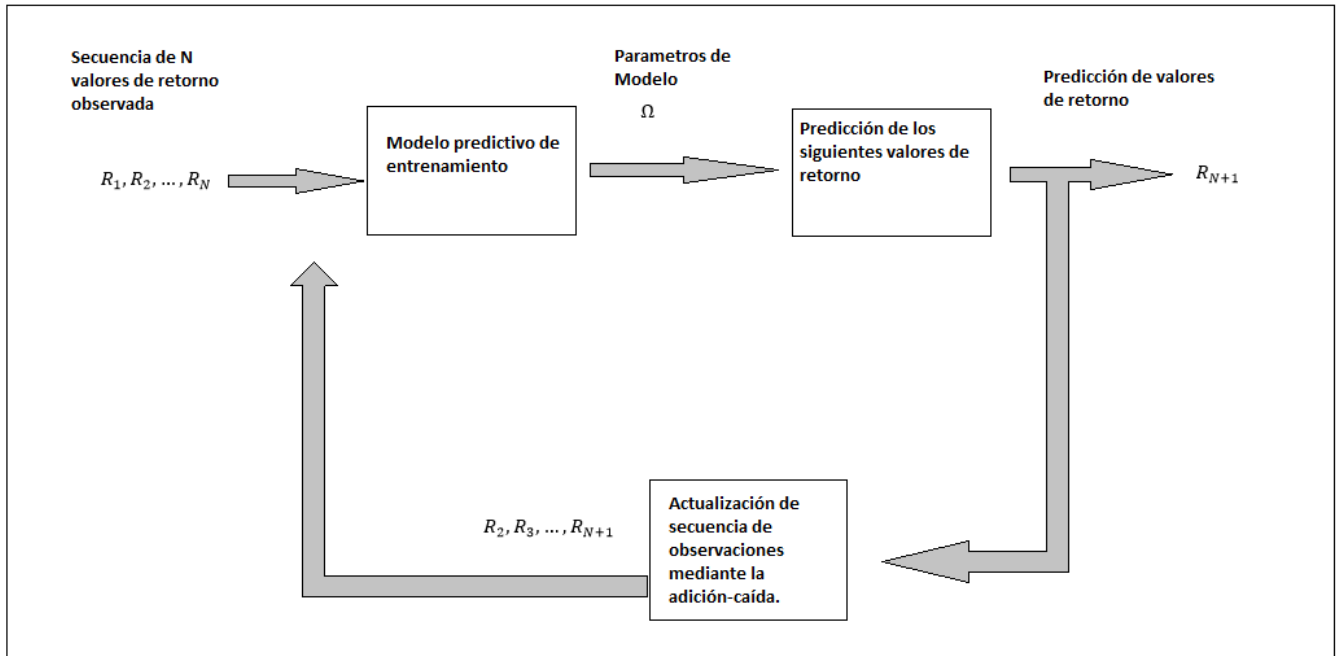


Figura 2 Modelo predictivo aplicado a series de tiempo en finanzas

## 2.1 Representación esparsiva, aprendizaje de diccionarios, bases y reconstrucción de series de tiempo

Series de tiempo o señales discretas en el tiempo  $\mathbf{x}$  son vectores en el espacio euclidiano  $n$ -dimensiones  $\mathbb{R}^n$ . El conjunto de señales atómicas  $\mathbf{D} = \{\mathbf{d}_i\}_{i=1}^k$  es llamado una base para el conjunto  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^m$  si las señales atómicas también llamados átomos son linealmente independientes y abarcan  $\mathbf{X}$ . Esto implica que cada señal en  $\mathbf{X}$  puede ser únicamente reconstruida por una sola combinación lineal de átomos en  $\mathbf{D}$ .

$$\mathbf{x} = \sum_{i=1}^k \alpha_i \mathbf{d}_i = \mathbf{D} \boldsymbol{\alpha} \quad (4)$$

Donde  $\mathbf{D} \in \mathbb{R}^{n \times k}$  y es una matriz compuesta de  $k$  columnas  $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_k]$  y las entradas del vector  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_k]^T$  son los coeficientes para esta combinación lineal. Si los elementos de la base  $\mathbf{D}$  son mutuamente ortonormales, es decir,  $\mathbf{d}_i^T \mathbf{d}_j = 0$  if  $i \neq j$  y cada átomo se encuentra en la hiper- esfera  $\mathbf{d}_i^T \mathbf{d}_i = 1 (\mathbf{D}^T \mathbf{D} = \mathbf{I})$ , entonces la base es completada tal como el caso de la base de Fourier.

Usar átomos linealmente independientes hace que el diccionario  $\mathbf{D}$  se convierte en un diccionario sobre-redundante existiendo múltiples opciones para el vector  $\boldsymbol{\alpha}$  para la reconstrucción de la señal  $\mathbf{x}$  de acuerdo con  $\mathbf{x} = \mathbf{D} \boldsymbol{\alpha}$ . Decimos que la señal  $\mathbf{x}$  admite una representación esparsiva sobre la base  $\mathbf{D}$  cuando se reconstruye de una forma concisa con pocos átomos y la señal es caracterizada por un factor esparsivo  $L$  cuando su representación esparsiva  $\boldsymbol{\alpha}$  tiene en la mayoría de las entradas de  $L$  diferente de cero (Rosas,2014).

$$\|\boldsymbol{\alpha}\|_0 \leq L \quad (5)$$

La representación resultante es un poderoso modelo para la reducción de almacenamiento y transmisión sugiriendo óptimos diccionarios sobre-redundantes para diferentes clases de señales.

## 2.2 Aprendizaje de diccionarios sobre-redundantes

En lugar de usar diccionarios predefinidos tales como wavelets para la reconstrucción esparsiva de señales los diccionarios se pueden ajustar a un conjunto de señales de entrenamiento. Dejando las columnas de la matriz  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m] \in \mathbb{R}^{n \times m}$  siendo el conjunto de  $m$  señales de entrenamiento de una clase particular para ser reconstruida esparsivamente a través de un óptimo diccionario sobre-redundante  $\mathbf{D}$ . La tarea del diccionario de aprendizaje es descrito por el siguiente problema de optimización.

$$[\mathbf{D}, \mathbf{A}] = \arg \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 \text{ subject to } \|\alpha_i\|_0 < L \forall i = 1, 2, \dots, m \quad (6)$$

Donde las columnas de la matriz  $\mathbf{A} = [\alpha_1, \alpha_2, \dots, \alpha_m] \in R^{k \times m}$  son las representaciones esparsivas del conjunto de señales en  $\mathbf{X}$  siendo  $L$  el factor esparsivo. Diferentes enfoques respecto al aprendizaje de diccionarios han sido desarrollados y basados en dos pasos:

- Encontrar una representación esparsiva  $\mathbf{A}$  de la señal de entrenamiento  $\mathbf{X}$  basado en un diccionario fijo  $\mathbf{D}$  a través de un algoritmo de búsqueda tales como Matching Pursuit (MP) (Mallat & Zhang, 1993) u Orthogonal Matching Pursuit (OMP) (Patti et al., 1993).
- Las señales atómicas son reconstruidas asumiendo coeficientes de reconstrucción fijos.

Un algoritmo adecuado para adaptar diccionarios en la representación esparsiva es el método K-SVD propuesto por Ahron et al. (2006) que expresa la reconstrucción total de la función de error.

$$\|\mathbf{X} - \mathbf{DA}\|_F^2 = \|\mathbf{X} - \sum_{i=1}^k \mathbf{d}_i \alpha_i^T\|_F^2 = \|\mathbf{X} - \sum_{i \neq j}^k \mathbf{d}_i \alpha_i^T - \mathbf{d}_j \alpha_j^T\|_F^2 = \|\mathbf{E}_j - \mathbf{d}_j \alpha_j^T\|_F^2, \quad (7)$$

Donde el átomo  $\mathbf{d}_i$  es la  $i$ th columna de  $\mathbf{D}$  y  $\alpha_i^T$  es el  $i$ th renglón de  $\mathbf{A}$ . K-SVD es un proceso iterativo de dos pasos para entrenar un diccionario.

El primer paso consiste en estimar la representación esparsiva de  $\mathbf{A}$  sujeto a:

$$\mathbf{A} = \arg \min_{\mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 \text{ sujeto a } \|\alpha_i\|_0 < L; \quad i=1, 2, \dots, m \quad (8)$$

En el segundo paso cada átomo  $\mathbf{d}_i$  correspondiente al renglón  $\alpha_i^T$  son encontrados utilizando SVD (Singular Value Decomposition) de acuerdo a:

$$[\mathbf{d}_j, \boldsymbol{\alpha}_j^T] = \arg \min_{\mathbf{d}_j, \boldsymbol{\alpha}_j^T} \|\mathbf{E}_j \boldsymbol{\Omega}_j - \mathbf{d}_j \boldsymbol{\alpha}_j^T\|_F^2, \quad (9)$$

donde la matriz  $\boldsymbol{\Omega}_j$  reduce  $\mathbf{E}_j$  manteniendo sólo las columnas que tienen coeficientes de reconstrucción diferente de cero en  $\boldsymbol{\alpha}_j^T$ . Mediante el uso del SVD la matriz  $\mathbf{E}_j \boldsymbol{\Omega}_j$  pueden ser aproximadas por una matriz de rango-1:

$$\mathbf{E}_j \boldsymbol{\Omega}_j = \sum_k \sigma_k \boldsymbol{\mu}_k \mathbf{v}_k^T \approx \sigma_1 \boldsymbol{\mu}_1 \mathbf{v}_1^T; \sigma_1 > \sigma_2 > \dots, \quad (10)$$

Donde  $\mathbf{d}_j = \boldsymbol{\mu}_1$  y  $\boldsymbol{\alpha}_j^T = \sigma_1 \mathbf{v}_1^T$ .