

# Construcción de un índice coincidente por medio de factores comunes dinámicos

Wilmer Osvaldo Martínez Rivera  
Departamento de Estadística  
Universidad Nacional de Colombia

Dirigido por:  
Fabio Humberto Nieto Sánchez  
PhD

Tesis para optar por el título de  
Maestría en Ciencias Estadística  
2010

# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Marco teórico</b>	<b>3</b>
2.1. Índice coincidente de Nieto y Melo . . . . .	3
2.2. Puntos de quiebre . . . . .	6
2.3. Propuesta de Banerji para el cálculo de índices líderes . . . . .	8
2.3.1. Número de periodos de adelanto del índice líder . . . . .	11
2.4. Análisis de factores comunes en series temporales multivariadas . . . . .	12
<b>3. Metodología propuesta para el cálculo del índice coincidente</b>	<b>15</b>
3.1. Perfil coincidente . . . . .	15
3.2. Ejemplificación de la metodología . . . . .	17
<b>4. Aplicaciones</b>	<b>23</b>
4.1. Sector macroeconómico . . . . .	23
4.2. Sector financiero . . . . .	30
4.3. Sector climatológico . . . . .	36
<b>5. Conclusiones y recomendaciones</b>	<b>45</b>
<b>A. Apéndice</b>	<b>50</b>

# Índice de figuras

2.1. Puntos de giro calculados por la metodología de <a href="#">Bry &amp; Boschan (1971)</a> . Los máximos están señalados con X y los mínimos con O . . . . .	9
3.1. Factor simulado (serie de referencia, línea negra), factor estimado con $\hat{\mathbf{P}}^0$ (línea negra punteada) y el factor estimado con $\hat{\mathbf{P}}$ (línea roja) . . . . .	19
3.2. Perfil principal para el factor simulado y factor estimado . . . . .	19
3.3. Perfil coincidente para procesos simulados con distintas varianzas, I hace referencia a la matriz identidad de orden 3 . . . . .	22
4.1. Series económicas concernientes a la presente aplicación . . . . .	24
4.2. Factores estimados y estimación del PIB mensual . . . . .	25
4.3. Perfiles coincidentes para cada factor. . . . .	28
4.4. Índice coincidente estimado con la metodología propuesta (línea roja) e Índice estimado por <a href="#">Castillo &amp; Nieto (2008)</a> (línea negra). . . . .	29
4.5. Perfil coincidente entre el índice estimado en el presente documento y el índice calculado por <a href="#">Castillo &amp; Nieto (2008)</a> . . . . .	29
4.6. Acciones de la BVC seleccionadas en la aplicación . . . . .	32
4.7. Continuación... Acciones de la BVC seleccionadas en la aplicación . . . . .	32
4.8. Factores estimados con las acciones seleccionados de la BVC . . . . .	33
4.9. Continuación... Factores estimados con las acciones seleccionados de la BVC e IGBC . . . . .	33
4.10. Factores estimados definitivos con las acciones seleccionados de la BVC e IGBC	34
4.11. Perfil coincidente aplicación acciones de la BVC. . . . .	34
4.12. Series climatológicas consideradas en la aplicación . . . . .	39
4.13. Factores estimados con los datos climatológicos en la Sabana de Bogotá . . . .	40

4.14. Factores comunes definitivos, estimados con los datos climatológicos en la Sabana de Bogotá e índice estimado por <a href="#">Rincón (2003)</a> . . . . .	41
4.15. Diagramas de Perfil coincidente para el estado del clima . . . . .	42

# Capítulo 1

## Introducción

Un índice puede considerarse como una predicción óptima de un proceso estocástico no observable con base en procesos observables. Por ejemplo, si se considera el estado de la economía, de las finanzas, del clima, entonces, algunos procesos observables de estos estados son: la producción, el empleo, el ingreso, el comercio, el consumo etc., en la actividad económica; los cambios de precio de las acciones inscritas en la bolsa de valores, en las finanzas; las fluctuaciones de las series de temperatura, humedad relativa, precipitación en la climatología. Así pues, un índice en el sector macroeconómico, financiero o climatológico, permite describir estas realidades.

En general, se construyen tres índices, a saber: el primero denominado índice “coincidente” el cual se refiere a la situación económica, financiera o climatológica contemporánea; el segundo conocido como índice “líder” o adelantado que alude a la situación futura y el tercero índice “rezagado”. Estos índices se construyen con base en los promedios ponderados de los procesos observables.

En los últimos años se han desarrollado diversas metodologías para construir índices coincidentes económicos, entre las cuales se encuentra la formulada por [Stock & Watson \(1989\)](#), que incluye el uso de modelos probabilísticos, la planteada por [Nieto & Melo \(2001\)](#) que consiste en una modificación de la propuesta de [Stock & Watson \(1989\)](#) y la desarrollada por [Castillo & Nieto \(2008\)](#) que estima el índice coincidente a partir de métodos bayesianos y el muestreador de Gibbs.

En términos prácticos, se encuentra que las variables utilizadas para el análisis económico, financiero o climatológico pueden ser explicadas por varios factores

---

comunes, en cuyo caso la pregunta natural es ¿cuál de estos factores representa o rastrea contemporáneamente el estado de la economía, el de las finanzas o el del clima?. Para responder esta pregunta (planteada en el documento de [Martínez & Nieto \(2007\)](#)), en el presente trabajo se desarrolla una metodología que permite determinar cuál de los factores comunes dinámicos, calculados por la metodología propuesta por [Peña & Poncela \(2006\)](#), es el que mejor sigue o representa el estado actual de la economía, de las finanzas o del clima, según el caso. El procedimiento se basa en el enfoque propuesto por [Banerji \(1999\)](#) para el cálculo de los índices líderes, en el cual, a partir de un conjunto de indicadores observados, permite predecir los puntos de giro del estado de la economía. Una vez se han evaluado estos indicadores, y se encuentra que adelantan o predicen los puntos de giro del estado de la economía, ellos se denominan índices candidatos a ser índices “líderes”. Finalmente, se evalúa la capacidad de cada uno de estos índices para adelantar los puntos de giro de la serie de referencia.

# Capítulo 2

## Marco teórico

### 2.1. Índice coincidente de Nieto y Melo

En esta sección se presenta una breve descripción del modelo estadístico de [Stock & Watson \(1989, 1991\)](#) (SW) utilizado para estimar el índice coincidente de la actividad económica, junto con la modificación realizada por [Nieto & Melo \(2001\)](#). El modelo de SW es una versión paramétrica de los modelos de ‘índices simples’ discutidos por [Sargent & Sims \(1977\)](#). Esta metodología se basa en modelos probabilísticos de estados que permiten estimar variables no observables, como el estado de la economía. [Nieto & Melo \(2001\)](#) modifican el modelo de estados desarrollado por SW. Algunos de los cambios buscan obtener la propiedad de estado de equilibrio del modelo, ya que el propuesto originalmente no la alcanza. Además, en él se permite no solo que las variables coincidentes estén cointegradas sino la estacionalidad dentro del modelo.

La metodología propuesta por [Nieto & Melo \(2001\)](#) supone que existe un grupo de  $n$  variables económicas observables e integradas de orden uno,  $X_{1t}, \dots, X_{nt}$ , que están relacionadas contemporáneamente con el estado de la economía y que además pueden tener otros movimientos que no están asociados con esta variable. Es decir, cada serie coincidente, además de tener un componente atribuible a la variable común, tiene un componente único o idiosincrático que no está correlacionado con los componentes idiosincráticos de las otras variables ni con el factor común. Así pues, en dicho modelo, las variables  $X_{it}$  tienen dos componentes estocásticos: la variable común no observada o estado de la

## 2.1 Índice coincidente de Nieto y Melo

---

economía, definida como un proceso estocástico latente denotado por  $\{C_t\}$ , y un componente  $u_{it}$  que representa los movimientos idiosincráticos de las series. De esta forma, la relación coincidente esta dada por la siguiente ecuación:

$$X_{it} = \beta_{it} + \gamma_i C_t + u_{it} \quad (2.1)$$

para todo  $t = 1, \dots, N$ , y para todo  $i = 1, \dots, n$ , siendo  $N$  la longitud del periodo muestral y  $n$  el número de variables coincidentes. Además,  $\beta_{it}$  es una componente determinística que puede incluir componentes estacionales,  $\gamma_i$  es una constante que representa la ponderación de  $C_t$  en  $X_{it}$  y  $u_{it}$  es la componente estocástica específica de  $X_{it}$ , independiente de  $C_t$ , y que obedece el proceso autoregresivo estacionario:

$$D_i(B)u_{it} = \epsilon_{it}, \quad (2.2)$$

donde  $D_i(B) = 1 - d_{i1}B - \dots - d_{ik}B^k$ , con  $B$  el operador de retraso y  $\epsilon_{it}$  ruido blanco gaussiano de media cero y varianza  $\sigma_i^2$ . Los procesos  $\{\epsilon_{it}\}$  se suponen mutuamente independientes, lo cual a su vez implica la independencia mutua de los procesos  $\{u_{it}\}$ .

La dinámica estocástica de  $\{C_t\}$  se describe por medio del siguiente modelo:

$$\phi(B)\Delta C_t = \delta + \eta_t, \quad (2.3)$$

siendo  $\phi(B)$  un operador estacionario autorregresivo de orden  $p$ ,  $\delta$  una constante y  $\{\eta_t\}$  un proceso ruido blanco con media cero y varianza  $\sigma_\eta^2$ . Esta ecuación refleja otro supuesto esencial en la metodología desarrollada y es que  $\{C_t\}$  es un proceso integrado de orden 1 [I(1)].

Si,  $\mathbf{X}_t = (X_{1t}, \dots, X_{nt})'$ ,  $\boldsymbol{\beta}_t = (\beta_{1t}, \dots, \beta_{nt})'$ ,  $\boldsymbol{\gamma}_t = (\gamma_{1t}, \dots, \gamma_{nt})'$ ,  $\mathbf{u}_t = (u_{1t}, \dots, u_{nt})'$ , y  $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{nt})'$ , entonces, las ecuaciones anteriores pueden ser reescritas de la siguiente forma vectorial:

$$\mathbf{X}_t = \boldsymbol{\beta}_t + \boldsymbol{\gamma}_t \mathbf{C}_t + \mathbf{u}_t, \quad (2.4)$$

$$\phi(B)\Delta \mathbf{C}_t = \delta + \eta_t, \quad (2.5)$$



$$D_i(B)\mathbf{U}_{it} = \epsilon_{it}. \quad (2.6)$$

Ahora bien, el objetivo es estimar (o predecir)  $\mathbf{C}_t$  con base en la información observada hasta el tiempo  $t$ , es decir calcular:

$$\mathbf{C}_{t|t} = \mathbf{E}[\mathbf{C}_t | \mathbf{X}_1, \dots, \mathbf{X}_t]$$

Para obtener dicha esperanza condicional, se puede usar el filtro de Kalman y en esa dirección las ecuaciones (2.4) - (2.6) deben ser transformadas en un modelo de estados. Para tal fin, sean  $C_{t+j|t} = E(C_{t+j} | C_0, C_1, \dots, C_t)$ ;  $j = 1, \dots, p$ ;  $\bar{C}_{t|t-1} = C_{t|t-1} - \delta$  y  $C_0 = 100$ .

Así, el modelo de estados se especifica mediante la ecuación del sistema o de transición (2.7) y la ecuación de observación o de medida por (2.8):

$$\boldsymbol{\alpha}_t = \boldsymbol{\mu}\delta + T\boldsymbol{\alpha}_{t-1} + R\boldsymbol{\zeta}_t \quad (2.7)$$

$$\mathbf{X}_t = \boldsymbol{\beta}_t + \mathbf{Z}\boldsymbol{\alpha}_t \quad (2.8)$$

donde  $\boldsymbol{\alpha}_t$  es el vector de estados del modelo y está especificado de la siguiente forma:

$$\boldsymbol{\alpha}_t = (C_t, C_{t+1|t}, C_{t+2|t}, \dots, C_{t+p|t}, \mathbf{u}'_t, \mathbf{u}'_{t-1}, \dots, \mathbf{u}'_{t-k})'. \quad (2.9)$$

Finalmente,  $\{C_{t|t}\}_t$  es el índice coincidente que con base en lo expuesto anteriormente esta dado de la siguiente manera:

$$C_{t|t} = C_0 + \sum_{j=1}^p \Delta C_{t+j|t} \quad (2.10)$$

donde  $\Delta C_{t+j|t} = E[\Delta C_{t+j} | X_1, \dots, X_t]$ . Más detalles en [Nieto & Melo \(2001\)](#).

## 2.2. Puntos de quiebre

El concepto de punto de quiebre está muy ligado a la evolución de la economía de un país. Se considera que los puntos de quiebre de una serie temporal son los puntos del tiempo en donde se presenta un máximo o mínimo local. Estos puntos intentan definir el comportamiento periódico de la serie y están asociados a los momentos de auge y recesión, propios de la economía. El segmento comprendido entre dos máximos (o dos mínimos) se denomina ciclo y el comprendido entre un máximo y un mínimo o viceversa se denomina fase.

Para identificar los puntos de quiebre en una serie temporal, [Bry & Boschan \(1971\)](#) proponen un procedimiento que tiene en cuenta algunos aspectos planteados por el National Bureau of Economic Research (NBER) de los Estados Unidos. El procedimiento se realiza sobre las series desestacionalizadas y consiste en aplicar sucesivamente un conjunto de reglas para identificar los puntos de giro en una secuencia de series filtradas, con diferentes grados de suavizamiento. Es decir, lo que se trata de buscar con los filtros es la componente de tendencia-ciclo de la serie. Para el cálculo de los puntos de giro, en este trabajo se utilizó el algoritmo de [Bry & Boschan \(1971\)](#) programado en RATS versión 7.1, el cual en resumen contempla los siguientes pasos:

- I** Los valores extremos de la serie son identificados y reemplazados por los obtenidos en una curva de Spencer<sup>1</sup>. Estos se definen como tales si la razón con respecto al valor obtenido al aplicar una curva de Spencer están fuera de 3.5 veces la desviación estándar de las razones.
  
- II** Después de haber corregido los datos extremos, en la serie original, se aplica un promedio móvil de orden 12 y sobre esta curva se identifican los puntos de giro así: primero se seleccionan los máximos y los mínimos. El valor de un mes<sup>2</sup> se puede catalogar como máximo si presenta el valor más alto entre los 5 meses precedentes y los 5 meses posteriores y de forma análoga se hace

---

<sup>1</sup>Una curva de Spencer es un promedio ponderado de 15 términos simétricos, donde las observaciones de los puntos extremos reciben pesos negativos y las observaciones del centro más peso y positivo.

<sup>2</sup>También se encuentra implementado para series con frecuencia trimestral

para un mes que se cataloga como mínimo. Luego se verifica que los puntos queden alternados, es decir, que un máximo sea seguido por un mínimo y viceversa. Si existen 2 o más máximos consecutivos se eliminan dejando el más alto y de la misma forma se procede para mínimos consecutivos.

**III** Se identifican los puntos de giro en la curva de Spencer. En principio se eliminan los vecinos, delimitados por  $\pm 5$  meses, de los puntos de giro determinados en el promedio móvil de orden 12. Con base en estos vecinos se seleccionan los giros semejantes sobre la curva de Spencer, es decir, los vecinos de los máximos del promedio móvil de orden 12, se investiga por los puntos más altos sobre la curva de Spencer y de igual manera se procede para los mínimos. Estos giros, identificados en la curva de Spencer, se someten a dos pruebas: una para garantizar que los puntos semejantes deben estar separados por al menos 15 meses y la otra para avalar que estos giros estén alternados. Así, si se excluye un giro se debe eliminar el opuesto para garantizar que queden alternados.

**IV** Los ciclos obtenidos en la curva de Spencer presentan ciertos inconvenientes con respecto a la serie original, pues esta curva puede cambiar giros, afectar pendientes y convertir fluctuaciones irregulares en patrones ondulatorios. Por esto se hace necesario aplicar, a los datos desestacionalizados, una media móvil de corto plazo denominada curva MCD<sup>1</sup> y sobre la media móvil de orden MCD se identifican los puntos de giro. Al igual que en el caso anterior un máximo se define como el valor más alto dentro de los 5 meses anteriores y posteriores al máximo correspondiente identificado en la curva de Spencer. De forma similar se procede para identificar los mínimos.

**V** El último paso de este procedimiento consiste en encontrar los máximos y los mínimos en los datos sin suavizar, de acuerdo a los giros establecidos en la curva MCD. Para ello se procede de manera análoga como se hizo al pasar del promedio móvil de orden 12 a la curva de Spencer y de la curva de

---

<sup>1</sup>El MCD (Month for Cyclical Dominance) de una serie, es el número de meses requerido para que la tendencia ciclo se afirme contra la componente irregular. El valor del MCD se puede obtener de tal forma que el promedio de cambio de la tendencia ciclo exceda al promedio de cambio en la componente irregular

## 2.3 Propuesta de Banerji para el cálculo de índices líderes

---

Spencer a la curva MCD, pero en esta parte el procedimiento establece un máximo, o un mínimo, como el valor más alto, o más bajo, en los datos sin suavizar dentro del máximo entre 4 y el valor del MCD.

No se aceptan giros si estos se encuentran en los últimos o primeros seis meses de la serie. De igual manera se establece que el primer y último máximo, o mínimo, debe ser al menos tan alto, o tan bajo, como algún valor entre estos y el extremo de la serie.

Un ejemplo del procedimiento descrito anteriormente es el siguiente: a la serie situación económica actual de la industria, la cual se detalla en la sección 4.1, se le aplica el procedimiento para calcular los puntos de quiebre, encontrando los resultados que se presentan en los Cuadros 2.1 y 2.2. En estos se hallan las fechas en las cuales se presentan los puntos de quiebre calculados en cada una de los pasos descritos anteriormente, en particular, en el Cuadro 2.2, se encuentran las fechas en que se presentan los puntos de giro definitivos, o del paso V. Además, en la Figura 2.1 se encuentran resaltados los puntos de quiebre, donde los máximos son indicados con (X) y los mínimos con (O), correspondientes a los puntos de giro calculados en el paso V; claramente, se trata de máximos y mínimos locales, así que el procedimiento es adecuado para el objetivo de este trabajo.

## 2.3. Propuesta de Banerji para el cálculo de índices líderes

Banerji (1999) propone una metodología para evaluar los indicadores candidatos a ser índices líderes, con base en la prueba de aleatorización para pares igualados de Fisher (ver Lehmann & Romano (2005)), esta prueba se utiliza porque no requiere supuestos acerca de la distribución del número de periodo líder del índice. Para ello se calculan las diferencias entre los puntos de giro, obtenidos por el algoritmo de Bry & Boschan (1971), de la serie de referencia y sus correspondientes en el indicador a evaluar, esto con el fin de determinar si los puntos de quiebre de cada indicador se presentan anticipadamente a los puntos de giro de la serie de referencia. Así pues, suponiendo que se observan  $N$  puntos

## 2.3 Propuesta de Banerji para el cálculo de índices líderes

Cuadro 2.1: Puntos de giro en los pasos I a IV

Pasos I y II		Paso III		Paso IV	
Máximos	Mínimos	Máximos	Mínimos	Máximos	Mínimos
	1982:01		1981:10		1981:11
1982:04	1983:03	1982:07	1983:04	1982:07	1983:04
1984:06	1985:03	1984:04	1984:12	1984:04	1984:12
1987:12	1989:04	1987:09	1989:09	1987:08	1989:10
1990:06	1991:06	1990:05	1991:09	1990:05	1991:09
1992:11	1993:12	1992:10	1993:09	1992:09	1994:01
1994:06	1996:09	1994:09	1996:10	1994:10	1996:10
1997:10	1996:10	1998:01	1999:04	1998:02	1999:04
2000:11	1999:02	2000:10	2001:06	2000:09	2001:06
2004:07	2001:10	2004:05	2005:02	2004:04	2005:02
	2004:11				

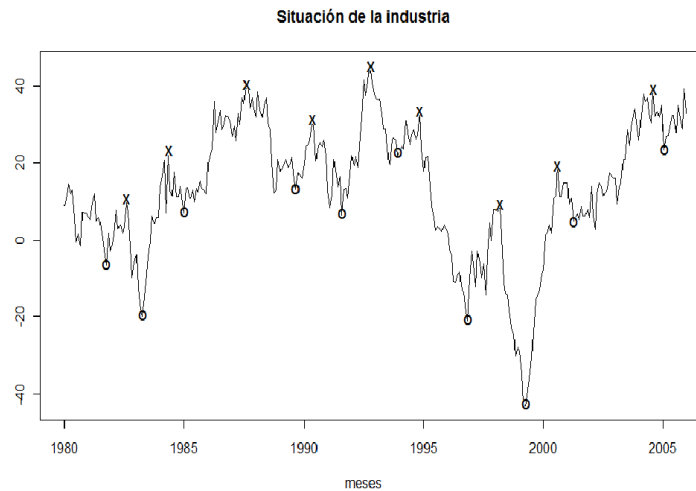


Figura 2.1: Puntos de giro calculados por la metodología de [Bry & Boschan \(1971\)](#). Los máximos están señalados con X y los mínimos con O

de quiebre, en cada una de las series, se obtienen  $N$  diferencias que pueden ser

### 2.3 Propuesta de Banerji para el cálculo de índices líderes

Cuadro 2.2: Puntos de giro para la variable situación económica, Paso V

Máximos	Máx a mín	Mín a máx	Mínimos	Mín a máx	Máx a mín
			1981:10		
1982:08		10	1983:04	18	8
1984:05	21	13	1985:01	21	8
1987:08	39	31	1989:09	56	25
1990:05	33	8	1991:08	23	15
1992:10	29	14	1993:12	28	14
1994:11	25	11	1996:11	35	24
1998:03	40	16	1999:04	29	13
2000:08	29	16	2001:04	24	8
2004:08	48	40	2005:01	45	5

Las distancias de las columnas 2, 3, 5 y 6 estan dadas en meses.

$\pm d_1, \pm d_2, \dots, \pm d_N$ , para un total de, a lo más,  $2^N$  posibles diferencias.

Como el objetivo de la propuesta de Banerji es determinar si los puntos de giro, de los indicadores candidatos a ser líderes, se presentan anticipadamente a los giros de la serie de referencia, él sugiere probar la siguiente hipótesis:

$$H_0 : \sum_{i=1}^N d_i = 0 \quad \text{contra la alternativa} \quad H_1 : \sum_{i=1}^N d_i > 0, \quad (2.11)$$

Para probar  $H_0$  se suman las diferencias observadas, luego de asignar signos negativos sistemáticamente a cada una de las diferencias  $d_i$ , y se calcula el número de sumas mayores o iguales a la suma observada de los  $2^N$  posibles resultados. Al número de resultados que cumplen la condición dada se denomina  $R$ , y así, se rechaza la hipótesis nula con un nivel de significancia de  $R/2^N$ . Lo anterior se aplica para las diferencias entre máximos y mínimos, individualmente.

Para ilustrar esta prueba supongamos que se encontraron  $N = 4$  máximos tanto para la serie de referencia como para el indicador. Se calculan las diferencias entre los máximos del indicador y la serie de referencia, las cuales son:

$$D = \{12, 4, -27, 5\},$$

## 2.3 Propuesta de Banerji para el cálculo de índices líderes

---

siendo su suma  $S = -6$ . Esta suma se compara con la obtenida después de asignar signos negativos de manera sistemática a cada  $d_i \in D$ . En el Cuadro 2.3 se presentan los  $2^4 = 16$  posibles resultados, allí se aprecia que 9 de estas sumas son mayores o iguales a la suma observada y por tanto el  $p$ -valor =  $9/16 = 0,5625$ . Así pues, no se rechaza  $H_0$ , es decir, no se encuentran diferencias significativas entre los puntos de giro del indicador en cuestión y la serie de referencia.

Cuadro 2.3: Ejemplo propuesta de Banerji (1999)

Asignación	Distancias $d_i$				Suma
1	12	4	-27	5	-6
2	-12	4	-27	5	-30
3	12	-4	-27	5	-14
4	-12	-4	-27	5	-38
5	12	4	27	5	48
6	-12	4	27	5	24
7	12	-4	27	5	40
8	-12	-4	27	5	16
9	12	4	-27	-5	-16
10	-12	4	-27	-5	-40
11	12	-4	-27	-5	-24
12	-12	-4	-27	-5	-48
13	12	4	27	-5	38
14	-12	4	27	-5	14
15	12	-4	27	-5	30
16	-12	-4	27	-5	6

Ahora bien, como el objetivo de la propuesta de Banerji (1999) es evaluar los indicadores a ser índices líderes, en el caso anterior si se rechaza  $H_0$ , la pregunta a contestar es: ¿cuántos periodos adelantan los puntos de giro del indicador a los giros de la serie de referencia?

### 2.3.1. Número de periodos de adelanto del índice líder

Si se desea conocer si el indicador adelanta los puntos de giro de la serie de referencia en por lo menos un mes, se resta 1 a cada una de las diferencias observadas, esto es,  $d_i - 1$  con  $i = 1, \dots, N$ . Con estas nuevas diferencias se realiza

## 2.4 Análisis de factores comunes en series temporales multivariadas

---

la prueba de hipótesis formulada en (2.11) y se determina el nivel de significancia con el que se rechaza la hipótesis de no capacidad de adelanto de los puntos de giro.

Continuando el procedimiento, se puede determinar si el número de periodos líder excede en  $2, 3, \dots, K$ , a los puntos de giro de la serie de referencia, con simplemente sustraer  $2, 3, \dots, K$  a las diferencias observadas respectivamente e ir realizando cada una de las pruebas respectivas. Con los niveles de significancia de cada una ellas se construye un gráfico de barras, el cual relaciona el nivel de confianza con el que se rechaza la hipótesis de no capacidad para adelantar un número de periodos, el cual está dado por  $100(1 - (R/2^N))\%$  contra el número de periodos líder. A este gráfico se le denomina diagrama de perfil principal.

## 2.4. Análisis de factores comunes en series temporales multivariadas

Peña & Poncela (2006) presentan un procedimiento para construir un modelo de factores comunes dinámicos para un proceso estocástico vectorial. Ellos asumen que es posible explicar el vector de variables de interés por medio de un número de variables latentes inferior al número de variables observables. Como procesos estocásticos que son, estos pueden ser no estacionarios. El modelo que utilizan es el siguiente: sea  $\mathbf{y}_t$  un vector  $m$ -dimensional de observaciones, generado por un conjunto de  $r$  factores comunes no observados (con  $r < m$ ), de tal manera que,

$$\mathbf{y}_t = \mathbf{P}\mathbf{f}_t + \mathbf{e}_t, \quad (2.12)$$

donde  $\mathbf{f}_t$  es un vector  $r$ -dimensional de factores comunes,  $\mathbf{P}$  es la matriz de pesos ( $m \times r$ ), y  $\mathbf{e}_t$  es un vector  $m$ -dimensional normalmente distribuido, con vector de medias cero y matriz de covarianzas diagonal de rango completo  $\Sigma_e$ . Así, toda la estructura dinámica común proviene de los factores comunes  $\mathbf{f}_t$ . Además se supone que el vector de factores comunes sigue un proceso  $VARMA(p, q)$  dado por

$$\Phi(B)\mathbf{f}_t = \mathbf{d} + \Theta(B)\mathbf{a}_t, \quad (2.13)$$



## 2.4 Análisis de factores comunes en series temporales multivariadas

---

donde  $\Phi(\mathbf{B}) = \mathbf{I} - \Phi\mathbf{B} - \dots - \Phi_p\mathbf{B}^p$  y  $\Theta(\mathbf{B}) = \mathbf{I} - \Theta\mathbf{B} - \dots - \Theta_q\mathbf{B}^q$  son polinomios matriciales de dimensión  $r \times r$ ,  $\mathbf{B}$  es el operador de rezago tal que  $\mathbf{B}\mathbf{y}_t = \mathbf{y}_{t-1}$ , las raíces de la ecuación  $|\Phi(\mathbf{B})| = 0$  pueden estar sobre o fuera del círculo unitario,  $\mathbf{d}$  es un vector de constantes de tamaño  $r \times 1$  y  $\mathbf{a}_t \sim N_r(\mathbf{0}, \Sigma_a)$ , con  $\Sigma_a$  la matriz de varianzas y covarianzas de rango completo. Además,  $\mathbf{E}(\mathbf{a}_t\mathbf{a}'_{t-h}) = \mathbf{0}$ ; para todo  $h \neq 0$ . Los componentes del vector de factores comunes,  $\mathbf{f}_t$ , pueden ser o no estacionarios, y se asumen las condiciones usuales de invertibilidad para el modelo *VARMA* (ver [Lutkepohl \(2005\)](#), p. 222). Nótese que si  $\mathbf{y}_t = (y_{1t}, \dots, y_{mt})'$ , entonces,  $y_{it} = \sum_{j=1}^r p_{ij}f_{jt} + e_{it}$ ; con  $i = 1, \dots, m$ .

Se define la matriz de covarianza muestral generalizada de rezago  $k$ , notada por  $\mathbf{C}_y(k)$ , así:

$$\mathbf{C}_y(k) = \frac{1}{T^{2d+d'}} \sum_{t=k+1}^T (\mathbf{y}_{t-k} - \bar{\mathbf{y}})(\mathbf{y}_t - \bar{\mathbf{y}})'; \quad k = 0, 1, \dots, m, \quad (2.14)$$

$$\text{donde } \bar{\mathbf{y}} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t.$$

El procedimiento propuesto por [Peña & Poncela \(2006\)](#) tiene dos objetivos: el primero es determinar el número de factores comunes y el segundo predecir los factores comunes con base en los vectores de series observables  $\mathbf{y}_t$ ,  $t = 1, \dots, T$ .

Para el pimero, se usa una prueba ji-cuadrado que consiste en lo siguiente: se construye la matriz  $\widehat{M}_1(k)$ ,  $k = 1, 2, \dots, K$  dada por

$$\widehat{M}_1(k) = \left[ \sum_{t=k+1}^T (\mathbf{y}_t\mathbf{y}'_t) \right]^{-1} \sum_{t=k+1}^T (\mathbf{y}_t\mathbf{y}'_{t-k}) \left[ \sum_{t=k+1}^T (\mathbf{y}_{t-k}\mathbf{y}'_{t-k}) \right]^{-1} \sum_{t=k+1}^T (\mathbf{y}_{t-k}\mathbf{y}'_t). \quad (2.15)$$

Luego, se calculan sus valores propios y se ordenan ascendentemente. En seguida, se calcula el estadístico de prueba

$$S_{m-r} = -(T - k) \sum_{j=1}^{m-r} \log(1 - \widehat{\lambda}_j), \quad (2.16)$$

## 2.4 Análisis de factores comunes en series temporales multivariadas

---

el cual tiene una distribución asintótica  $\chi_{(m-r)^2}^2$  para todo  $k = 1, \dots, K$  bajo la hipótesis nula  $H_0$  : “hay máximo  $r$  factores comunes”;  $m$  es el orden de la matriz  $\widehat{M}_1(k)$  y  $\widehat{\lambda}_j$  sus valores propios. Así pues, dado  $k$ , e incrementando  $r$  en (2.16), el estadístico de prueba  $S$  se compara con el percentil  $(1 - \alpha)$  de la  $\chi_{(m-r)^2}^2$ . Si el estadístico de prueba es mayor que este percentil, para un cierto nivel de significancia  $\alpha$ , entonces, se rechaza la hipótesis nula de que halla un máximo de  $r$  factores comunes .

Para el segundo, una estimación inicial de la matriz de pesos  $\widehat{\mathbf{P}}^0$  que podrían ser los primeros vectores propios de la matriz de covarianzas generalizada, para el rezago  $k = 1$ , y a partir de esto una estimación inicial razonable para los factores es  $\widehat{\mathbf{f}}^0 = (\widehat{\mathbf{P}}^0)' \mathbf{y}_t$ .

La estimación del modelo de factores comunes, por máxima verosimilitud, se puede hacer vía el algoritmo EM de [Dempster \*et al.\* \(1977\)](#), tal como lo indica [Peña & Poncela \(2006\)](#).

# Capítulo 3

## Metodología propuesta para el cálculo del índice coincidente

### 3.1. Perfil coincidente

La prueba empleada por [Banerji \(1999\)](#), en la construcción del diagrama de perfil principal, es conocida como una prueba de permutación con alternativa a una cola. La construcción del diagrama de perfil principal se basa en el hecho de rechazar la hipótesis nula en (2.11), dado que el fin es encontrar diferencias esencialmente negativas o positivas dependiendo de cómo se planteen las diferencias  $d_i$ 's de los puntos de quiebre. Las hipótesis nulas, a probar en la construcción del perfil principal propuesto por [Banerji \(1999\)](#), se pueden representar como

$$H_0^{(l)} : \sum_{i=1}^N (d_i - l) = 0; \quad l = 0, 1, \dots, m \quad (3.1)$$

donde el valor de  $l > 0$ , para el cual no se rechaza la hipótesis nula, se considera el periodo líder.

De otra parte, como el objetivo de este trabajo es diseñar una metodología que permita calcular índices coincidentes o contemporáneos, para cumplir con él se tomó como base la propuesta de [Banerji \(1999\)](#), en el sentido de no rechazar  $H_0$  en (3.1), cuando  $l = 0$ . Además, interesa ver que pasa con  $H_0$ , para valores de  $l$  alrededor de 0, esto es, para  $l = \pm 1, \pm 2, \dots$  con alternativa a dos colas.

Bajo estas consideraciones, se construye un diagrama de barras, en el cual la altura de cada una de ellas es el  $p$ -valor (cuyo cálculo se presenta más en seguida) de la prueba (3.1) y la base es el valor de  $l$ . Así se obtiene lo que se ha denominado diagrama de perfil coincidente.

Dado que en la construcción de perfil coincidente se usan los  $p$ -valores de las pruebas en (3.1) con  $l = 0, \pm 1, \pm 2, \dots$  contra la hipótesis alternativa de dos colas, en cada caso, el diagrama de perfil principal propuesto por Banerji (1999) resulta ser un caso particular de lo expuesto en estas líneas.

Así pues, la metodología propuesta permite determinar, a partir de un conjunto de factores comunes estimados, cual de ellos sigue o rastrea mejor una serie de referencia. Aquel que cumpla esta condición se considera índice coincidente (de la economía, las finanzas o del estado del clima, según el caso). El factor común que se considera como índice coincidente es aquel cuyo diagrama de perfil coincidente, para  $l = 0$  en (3.1), no rechaza  $H_0$ , para un nivel de significancia establecido. Además, el  $p$ -valor, en este punto, es el mayor con respecto a los demás factores, en el caso de que halla dos o más.

#### Prueba de permutación

La prueba de permutación para dos muestras independientes, se basa en el rechazo de la hipótesis nula, en favor de la alternativa específica, si el nivel de significancia predeterminado es mayor que el  $p$ -valor correspondiente. El procedimiento para obtener un  $p$ -valor aproximado es equivalente a la prueba de permutación, basada en la diferencia de medias como se describe a continuación:

1. Se calcula la suma de las diferencias de los puntos de quiebre de cada factor con la serie de referencia, es decir  $D_0 = \sum_{i=1}^N d_i$ , donde los  $d_i$ 's son como se definen en la sección 2.3.
2. Se asignan signos a las diferencias observadas, esto es,  $\pm d_1, \pm d_2, \dots, \pm d_N$  para a lo más un total de  $2^N$  posibles aleatorizaciones (muestras).
3. Se calcula la suma  $D_i$  para cada una de las posibles muestras de los  $d_i$ 's, del paso anterior.

4. Para la prueba bilateral el  $p$ -valor está dado por

$$P_0 = \frac{\#\{D_i : |D_i| \geq |D_0|; \quad i = 1, 2, \dots, 2^N\}}{2^N}, \quad (3.2)$$

donde  $\#\{D_i : |D_i| \geq |D_0|; \quad i = 1, 2, \dots, 2^N\}$  denota el cardinal del conjunto de las posibles muestras que cumplen lo dicho en el numeral anterior, ver [Bautista & Gómez \(2007\)](#) y [Welch & Gutiérrez \(1988\)](#).

### 3.2. Ejemplificación de la metodología

Para ilustrar la metodología propuesta, en primer lugar se simuló un proceso  $\{f_t\}$ ,  $ARIMA(2, 1, 0)$  de 313 datos. Los coeficientes de la parte autorregresiva se tomaron de tal manera que las raíces del polinomio autorregresivo  $(1 - \phi_1 z - \phi_2 z^2)$  sean complejas, garantizando así ciclos en el proceso. Los coeficientes considerados son:  $\phi_1 = 1$  y  $\phi_2 = -0.6$  y un proceso de ruido blanco gaussiano  $\{\varepsilon_t\}$  con media 0 y varianza 1. Así pues, a partir del proceso simulado  $f_t = f_{t-1} - 0.6f_{t-2} + \varepsilon_t$ , se obtienen cada una de las componentes del vector  $\mathbf{y}_t$  como combinación lineal del proceso  $\{f_t\}$  más un proceso de ruido blanco gaussiano  $\{\mathbf{e}_t\}$  con vector de media  $\mathbf{0}$  y matriz de covarianzas  $\Sigma_e = \sigma^2 I$ ;  $\sigma^2 = 1$  como sigue:

$$\mathbf{y}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \\ 2 \end{pmatrix} f_t + \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix} \quad (3.3)$$

Ahora bien, haciendo uso de la metodología planteada por [Peña & Poncela \(2006\)](#), se calcula  $\hat{f}_t$ , luego se obtienen los puntos de quiebre tanto para  $\hat{f}_t$  como para  $f_t$  (serie de referencia) y se calculan las diferencias entre ellos. Siguiendo los pasos propuestos por dicha metodología (sección 2.4), se identifica el número de factores comunes y se obtiene una estimación inicial para la matriz de pesos y el factor común. Al aplicar la prueba (2.16) para  $k = 1, 2, 3$  se obtienen los resultados presentados en el Cuadro 3.1. Las entradas del Cuadro 3.1 son los valores del estadístico  $S$ , para cada rezago  $k$ , dichos valores son usados para probar un máximo de  $r$  factores comunes y deben ser comparados con una distribución

### 3.2 Ejemplificación de la metodología

$\chi^2$  con  $(m - r)^2$  grados de libertad. Para todos los rezagos se rechaza la hipótesis nula de 0 factores comunes, como era de esperarse, mientras que ella, no se rechaza para 1 factor común, en todos los rezagos.

Cuadro 3.1: Determinación del número de factores comunes para los datos simulados

r	rezago k		
	1	2	3
0	1348*	971*	786*
1	1.024	0.195	0.308
2	0	0	0

\* Indica que la hipótesis nula de máximo  $r$  factores es rechazada con un nivel de significación del 5%.

Para obtener una estimación inicial de la matriz de pesos  $\mathbf{P}$  y del factor común se construye la matriz de covarianzas generalizada. El primer vector propio asociado al mayor valor propio para el rezago  $k = 1$ , de la matriz  $Cy(k)$  en (2.14), constituye la estimación inicial de  $\mathbf{P}^0$ , denotado  $(\widehat{\mathbf{P}}^0)' = (0.44, 0.22, 0.87)$ . Con esta matriz se calcula una primera estimación del factor común, así  $(\widehat{\mathbf{f}}^0) = (\widehat{\mathbf{P}}^0)'(\mathbf{y}_t)$ . Ahora bien, para obtener la estimación final de la matriz de pesos y del factor común, se aplica el algoritmo **EM** asumiendo como condición inicial  $(\widehat{\mathbf{P}}^0)$  para la matriz de pesos del factor. De este modo se obtiene la estimación final de la matriz de pesos del factor  $(\widehat{\mathbf{P}}) = (0.473, 1.471, 0.472)$  y la estimación del factor (ver Figura 3.1).

Como el objetivo del ejercicio es comparar los puntos de giro del factor simulado con los del factor estimado (línea roja de la Figura 3.1), con el fin de evaluar la metodología planteada, en el Cuadro 3.2 se presentan los puntos en el tiempo para los cuales se tienen máximos y mínimos (puntos de giro) para cada uno de los factores, además las diferencias  $d_i$  entre los puntos de quiebre del factor simulado con los del factor estimado. A partir de estas diferencias se construye el perfil coincidente, el cual se presenta en la Figura 3.2. Allí se observa que el mayor  $p$ -valor se presenta en cero, que es lo que se esperaba. La línea roja representa el nivel de significación de la prueba al 5%.

## 3.2 Ejemplificación de la metodología

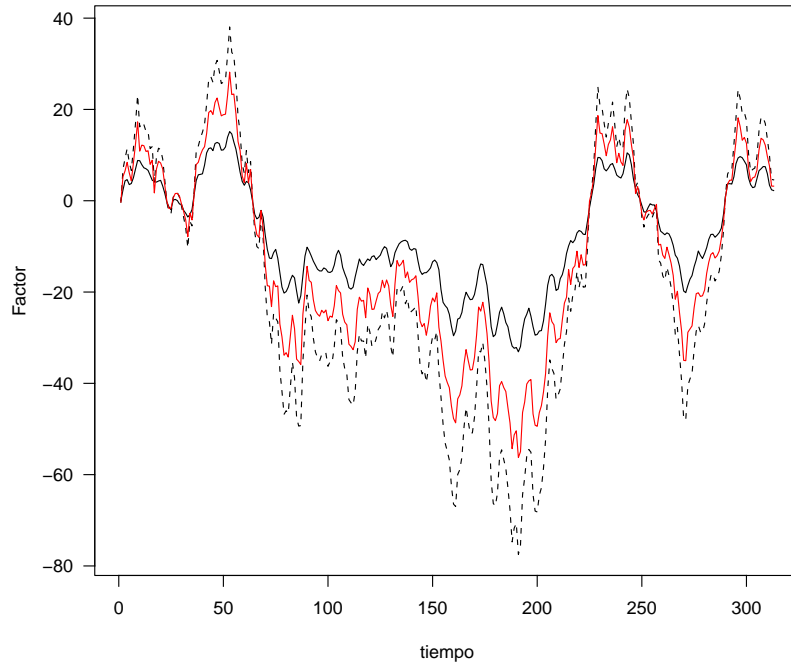


Figura 3.1: Factor simulado (serie de referencia, línea negra), factor estimado con  $\hat{\mathbf{P}}^0$  (línea negra punteada) y el factor estimado con  $\hat{\mathbf{P}}$  (línea roja)

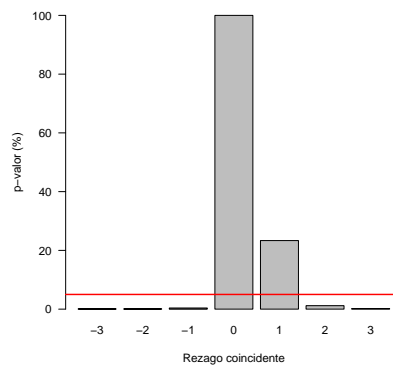


Figura 3.2: Perfil principal para el factor simulado y factor estimado

### 3.2 Ejemplificación de la metodología

---

Cuadro 3.2: Puntos de giro para el factor simulado y el factor estimado

Punto de giro	Factor simulado	Factor estimado	$d_i$
min	9	9	0
max	33	33	0
min	53	53	0
max	86	81	5
min	137	137	0
max	160	161	-1
min	173	174	-1
max	191	191	0
min	229	229	0
max	271	271	0
min	296	296	0

La razón del ejercicio anterior es la siguiente: se simuló  $\{f_t\}$  que representa el proceso de interés, el cual se supone no observable y cuyo objetivo es estimar. Para estimar  $\{f_t\}$  se hace uso del modelo de factores comunes dinámicos, propuesto por Peña & Poncela (2006). En este modelo (ver expresión 2.12) se asume que el proceso  $\{y_t\}$ , el cual es observado, proviene de un proceso latente  $\{f_t\}$  denominado factor común, así la metodología propuesta por Peña & Poncela (2006), expuesta previamente en la Sección 2.4 del Capítulo 2, ofrece una manera para estimar el proceso  $\{f_t\}$  en el modelo. Bajo esta idea se calculó  $y_t$ , como una combinación lineal del proceso simulado  $\{f_t\}$ , como se aprecia en (3.2). Por tanto se calculó  $\hat{f}_t$  y sus puntos de giro se compararon con los de  $f_t$ .

En los ejercicios con datos reales del sector económico, financiero y climatológico que se presentan en el siguiente Capítulo, la serie, que hará las veces del proceso latente  $\{f_t\}$ , denominada serie de referencia, se escoge en concordancia con la experiencia o con lo sugerido por los expertos en cada caso.

Puesto que cada una de las tres componentes del proceso  $\{y_t\}$  en (3.2) fueron simuladas como una combinación lineal de  $f_t$  más un error cuya varianza es  $\sigma^2 = 1$ , ahora se simulan 6 procesos  $\{y_t\}$  donde en cada caso la varianza del proceso de error  $\{e_t\}$  es  $\sigma^2 I$ , con  $I$  la idéntica y  $\sigma^2 = 0.09, 0.36, 1, 2.25, 4$  y  $25$ , es decir, en cada caso las componentes del vector  $y_t$  se simulan con el mismo error. Esto con el fin de observar qué sucede con los puntos de giro del factor estimado



## 3.2 Ejemplificación de la metodología

---

$\hat{f}_t$  respecto a los puntos de giro del factor simulado  $f_t$ , a medida que aumenta la varianza del error con la cual son simuladas las componentes del proceso  $\{\mathbf{y}_t\}$ .

Una vez se simula cada uno de los seis procesos  $\{\mathbf{y}_t\}$ , por medio de la metodología de Peña & Poncela (2006), en cada caso se calcula  $\hat{f}_t$  y se procede como en el primer ejercicio. A cada uno de los factores estimados se le calculan los puntos de quiebre y se comparan con los puntos de quiebre del factor simulado  $f_t$ , con las diferencias obtenidas entre los puntos de quiebre, en cada caso, se contruye el diagrama de perfil coincidente como se aprecia en la Figura 3.3. Allí se observan seis gráficos de perfil coincidente, cada uno de ellos correspondiente a las diferencias entre los puntos de giro del factor simulado y el factor estimado del proceso  $\{\mathbf{y}_t\}$ , simulado como se indicó anteriormente. Por ejemplo, la gráfica a) corresponde a las diferencias calculadas entre el factor simulado y el factor estimado del proceso  $\{\mathbf{y}_t\}$  simulado con un proceso de ruido cuya varianza es  $0,09I$ . En cada gráfica se indica la varianza del ruido con la cual fue simulado el proceso  $\{\mathbf{y}_t\}$ .

En estas gráficas se aprecia que a medida que aumenta la varianza del error, es decir a medida que se adiciona ruido al proceso  $\{\mathbf{y}_t\}$ , aumentan las diferencias entre los puntos de giro del factor estimado y el factor simulado y por ende disminuye la coincidencia entre los puntos de giro comparados. Por lo anterior, entre mayor sea la perturbación en el vector de observaciones mayor dificultad se tiene para extraer los factores comunes del proceso.

### 3.2 Ejemplificación de la metodología

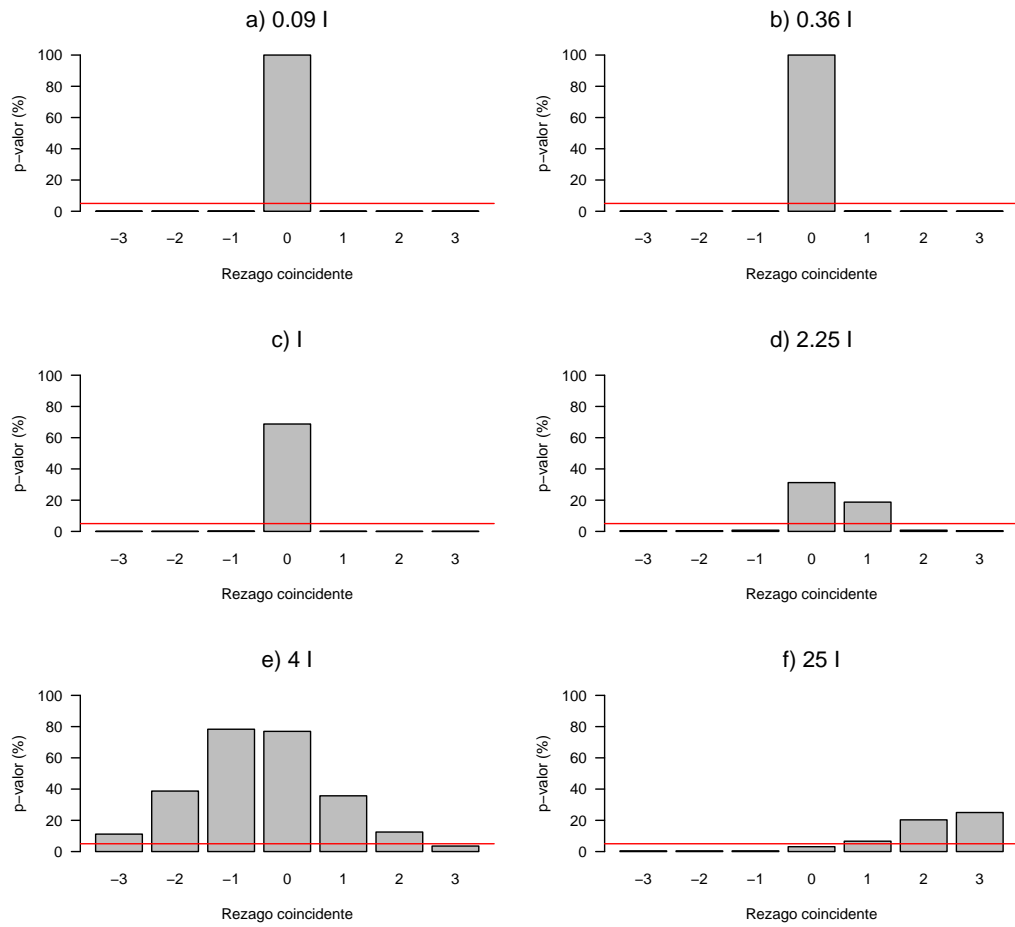


Figura 3.3: Perfil coincidente para procesos simulados con distintas varianzas, I hace referencia a la matriz identidad de orden 3

# Capítulo 4

## Aplicaciones

En el presente capítulo se llevan a cabo tres aplicaciones de la metodología propuesta, una en el sector macroeconómico, otra en el sector de las finanzas y una tercera en climatología; todas ellas con datos reales registrados en Colombia.

### 4.1. Sector macroeconómico

En Colombia se han hecho varios intentos por construir índices coincidentes del sector macroeconómico del país, como se aprecia en los trabajos de [Nieto & Melo \(2001\)](#), [Nieto \(2003\)](#), [Martínez & Nieto \(2007\)](#), entre otros. En estos trabajos se han encontrado diversos índices y en algunos casos han sido comparados con el PIB o el IPR de forma descriptiva, pero la pregunta es: ¿cuál de estos índices sigue o rastrea mejor el estado de la economía? Para responder tal interrogante, en esta aplicación se analizan las series seleccionadas en el documento de [Nieto & Melo \(2001\)](#) que son: situación económica actual de la industria, volumen actual de pedidos por atender de la industria, índice de producción real de la industria manufacturera sin trilla de café, índice de empleo de obreros de la industria, producción de cemento, demanda de energía más consumo de gas residencial e industrial, importaciones reales exceptuando las de bienes de capital y duraderos, cartera neta real de moneda legal y saldo de efectivo en términos reales. En [Nieto & Melo \(2001\)](#) se analizaron estas series desde enero de 1980 hasta agosto de 2001; pero en esta aplicación el periodo de observación se extiende hasta diciembre de 2005. Los datos se desestacionalizaron usando el método de media móvil aditivo

## 4.1 Sector macroeconómico

de EViews y fueron ajustados por intervenciones y datos atípicos (ver [Martínez & Nieto \(2007\)](#)). En la Figura 4.1 se presentan las nueve series mencionadas y ajustadas.

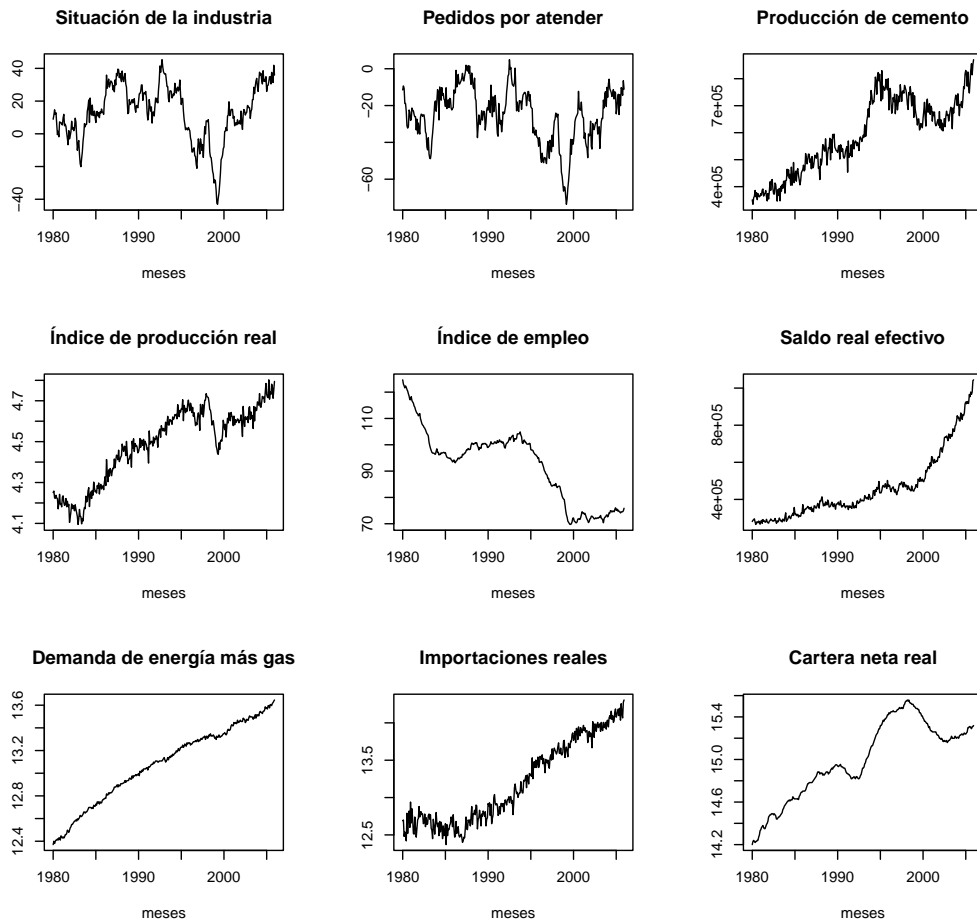


Figura 4.1: Series económicas concernientes a la presente aplicación

A este conjunto de datos se le aplicó lo propuesto por [Peña & Poncela \(2006\)](#), encontrando dos factores comunes. En el Cuadro 4.1 se encuentran los resultados de la prueba del número de factores y en el Cuadro A.1 del Apéndice A, se presenta la estimación de la matriz de pesos  $\mathbf{P}$  de los factores; en la Figura 4.2 se aprecia la estimación de los factores.

En este caso, como lo recomiendan los expertos, se consideró como serie

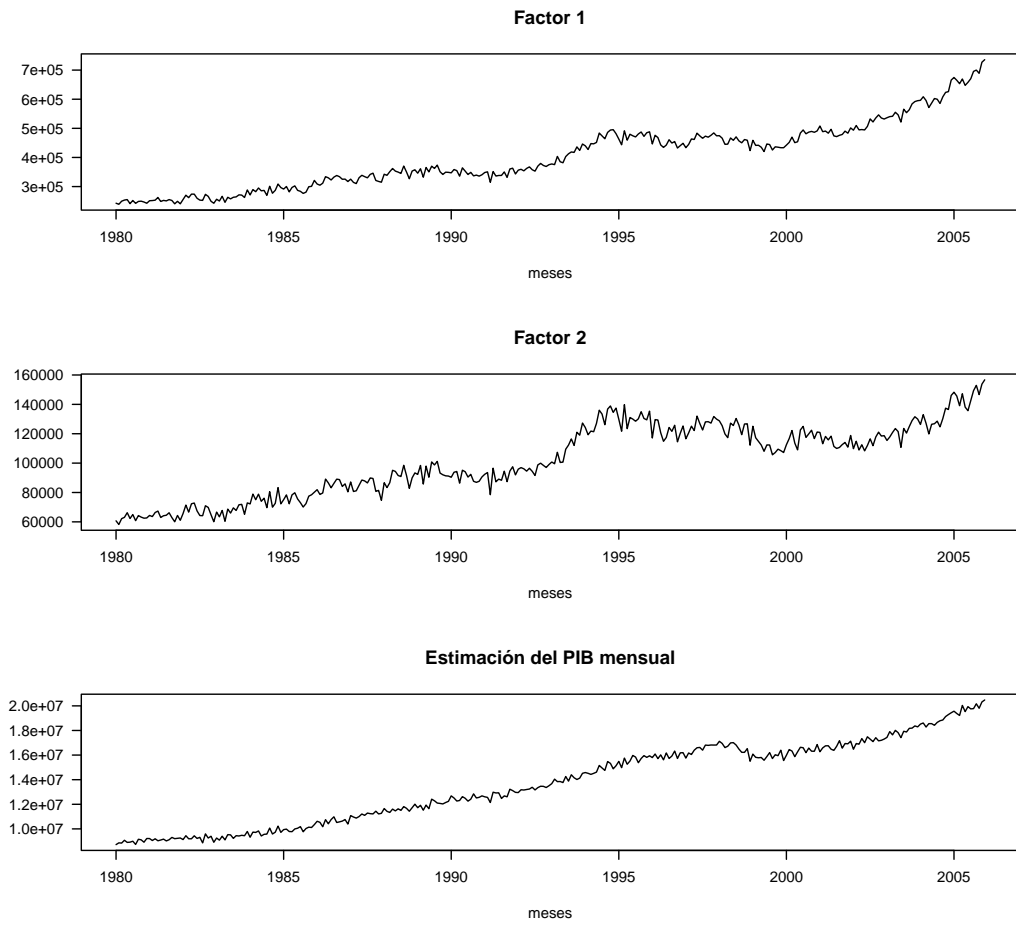


Figura 4.2: Factores estimados y estimación del PIB mensual

Cuadro 4.1: Determinación del número de factores, para los datos de la macroeconomía colombiana

r	rezago k				
	1	2	3	4	5
0	3006.2 *	2871.7 *	2833.7 *	2698.0 *	2655.1 *
1	794.8 *	773.8 *	762.4 *	703.8 *	704.4 *
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0

\* Indica que la hipótesis nula de máximo  $r$  factores es rechazada con un nivel de significancia del 5%.

de referencia el PIB, pero como esta serie en el país se mide trimestralmente y la frecuencia del conjunto de datos a analizar es mensual, se hace necesario desagregar la serie del PIB trimestral a mensual para lograr la comparación entre los puntos de giro. Para ello se utiliza la metodología de Nieto (2004).

En la aplicación empírica del trabajo de Nieto (2004) se hace una desagregación del PIB anual al PIB trimestral. Para ello se utiliza el conjunto de indicadores siguiente: producción de cemento, índice de producción de la industria sin trilla de café, consumo de energía más gas y producción de gasolina en miles de galones. Estos datos se tomaron del Banco de la República, para más detalles ver (Nieto (2004)).

Ahora bien, para cumplir con el objetivo de este ejercicio, se considera el mismo conjunto de indicadores en frecuencia mensual en el periodo comprendido entre enero de 1980 y diciembre de 2005 y se obtiene como resultado la estimación del PIB mensual como se aprecia en la gráfica inferior de la Figura 4.2.

Luego, se calculan los puntos de giro para cada uno de los factores y para el PIB, por medio del procedimiento de Bry & Boschan (1971). Pero desafortunadamente para el PIB solo se identifican dos puntos de giro, primero y último mes del año 1998, lo cual no permite apreciar los hechos estilizados de

la economía dado que la estimación del PIB es una curva muy suave como bien se aprecia en la gráfica de este. Ante este inconveniente y por sugerencia de los expertos se decide tomar como serie de referencia la serie del IPR, dado que dicha serie sigue los hechos estilizados más importantes de la economía colombiana. Por tal razón y continuando con la presente aplicación, en el Cuadro 4.2 se relacionan las fechas en las cuales se presentan máximos y mínimos locales (puntos de giro) del IPR y de los factores, en este Cuadro se presentan también las diferencias  $d_i$  entre los puntos de giro del IPR y cada factor. Dado que se encuentran diferencias hasta de 14 meses entre los puntos de quiebre, estas no permiten establecer claramente la cercanía entre los demás puntos de giro de los factores al IPR, por tanto se tomó como criterio considerar únicamente aquellas distancias menores o iguales a 4 meses, para la construcción del perfil coincidente.

Cuadro 4.2: Puntos de giro para los Factores comunes y el IPR

Pto de giro	IPR	Factor 1	Factor 2	$d_1$	$d_2$
max		abr-82			
min	dic-82	dic-82	oct-81	0	14
max	mar-90	ago-89	ago-89	7	7
min	mar-91	mar-91	mar-91	0	0
max	nov-95	nov-94	mar-95	12	8
min	mar-97	oct-96	oct-96	5	5
max	dic-97	nov-97	nov-97	1	1
min	may-99	ago-99	ago-99	-3	-3
max	nov-00	ene-01	jul-00	-2	4
min	may-02	jul-01	may-02	10	0

En la Figura 4.3 se presentan los diagramas de perfil coincidente para cada factor respecto al IPR. En estos diagramas se observa lo siguiente: en primer lugar, se aprecia que para los dos factores en el rezago cero, no se rechaza la hipótesis nula “la suma de las diferencias es cero”, para un nivel de significancia del 5%; pero para el factor 1 el  $p$ -valor, en este punto, es mayor que el  $p$ -valor del factor 2; en segundo lugar, el  $p$ -valor en el rezago cero, del perfil coincidente para el factor 1, es el mayor respecto a los demás rezagos. Dado lo anterior, se concluye que el factor común que mejor rastrea los hechos estilizados de la

economía colombiana es el factor 1 y por lo tanto, este factor se puede interpretar como un índice coincidente para el estado de la economía colombiana.

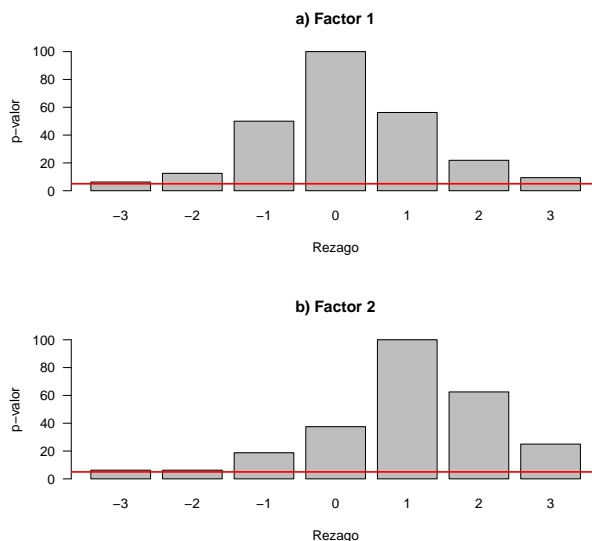


Figura 4.3: Perfiles coincidentes para cada factor.

Como [Castillo & Nieto \(2008\)](#) estimaron un índice del estado de la economía, el cual en general concuerda con los hechos estilizados de la economía, junto a este se presenta el índice estimado en el presente documento. Para ello se estandarizaron estos índices en el siguiente sentido: cada observación se corrige por la media de la variable y luego se divide por la desviación estándar, esto para cada serie, como se aprecia en la [Figura 4.4](#). Allí se observa la contracción en el periodo 1989 - 1991, además durante 1996 se presenta una caída de la actividad económica y finalmente, en el periodo comprendido entre 1998 y 1999 se presenta la mayor contracción, que en el índice estimado no es tan pronunciada.

A partir del resultado anterior, se concluye que tanto la metodología planteada por [Castillo & Nieto \(2008\)](#) como la propuesta por [Peña & Poncela \(2006\)](#) dan como resultado índices coincidentes similares, a la luz de la metodología propuesta, ver [Figuras 4.4 y 4.5](#). Lo anterior es importante dado que no solo se encontró coincidencia del índice estimado con el IPR, sino también con un índice estimado por otra metodología. Ahora bien, la ventaja de proceder con



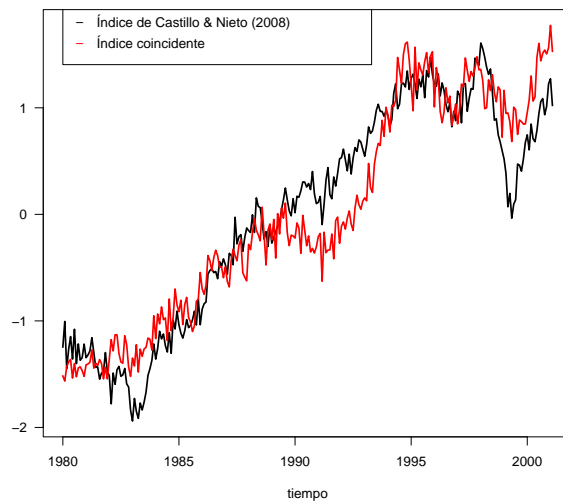


Figura 4.4: Índice coincidente estimado con la metodología propuesta (línea roja) e Índice estimado por [Castillo & Nieto \(2008\)](#) (línea negra).

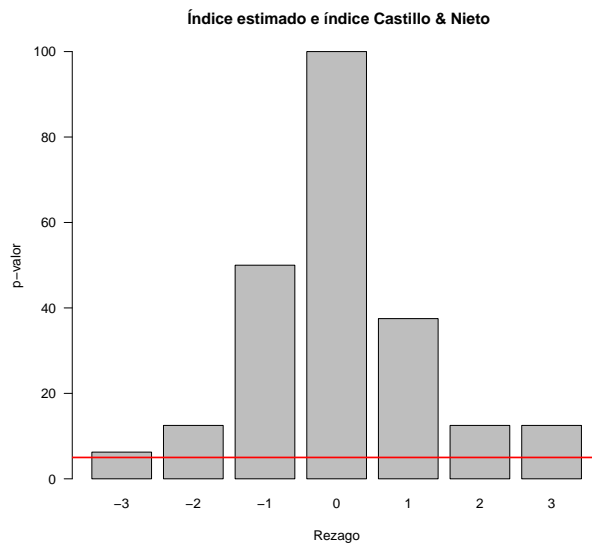


Figura 4.5: Perfil coincidente entre el índice estimado en el presente documento y el índice calculado por [Castillo & Nieto \(2008\)](#)

la metodología propuesta, es que como se hace uso de lo propuesto por [Peña & Poncela \(2006\)](#), respecto a lo propuesto por [Castillo & Nieto \(2008\)](#) es que se evalúa más de un factor, lo cual puede dar indicios de otro tipo de comportamiento no observable propio de la economía. Por tanto en términos prácticos se sugiere utilizar la metodología empleada en este trabajo, siendo además una metodología consistente, como se aprecia en los ejercicios de simulación de la sección 3.2 del Capítulo 3.

### 4.2. Sector financiero

Actualmente, en la Bolsa de Valores de Colombia (BVC) existen alrededor de cincuenta y dos empresas que tienen inscritas sus acciones. La idea es que el propietario de una empresa o acción tenga la posibilidad de ofrecer en el mercado público de valores sus acciones, de manera que las pueda vender a un mejor precio.

Según se presenta en la página web de la Bolsa de Valores de Colombia, BVC, una empresa emisora inscrita en ella está obligada a proporcionar información periódica y eventual, de tal manera que el mercado disponga de información suficiente y oportuna para la toma de decisiones. Esto garantiza transparencia, seguridad y liquidez en el mercado público de valores. Con la información disponible de las empresas inscritas en la BVC se construyen índices, los cuales constituyen la base para identificar la percepción del mercado frente al comportamiento de las empresas y de la economía, entre otros usos. Los criterios básicos para la construcción de un índice se definen en dos pasos: el primero, seleccionar la canasta de acciones o conjunto de acciones que pertenecen al índice y el segundo, determinar la manera en que se va a ponderar cada una de estas acciones seleccionadas dentro del índice (tomado de la página web [BVC](#)).

Para efectos de la presente aplicación, se toma como serie de referencia el Índice General de la Bolsa de Valores de Colombia (IGBC), este índice es el resultado de ponderar las acciones más líquidas y de mayor capitalización que se negocian en la Bolsa, es decir, aquellas que tienen una mayor rotación y frecuencia. Aquellas acciones que tengan una rotación mayor a 0.5 % semestral y una frecuencia superior a 40 % en el trimestre anterior, son incluidas dentro del IGBC. Este índice opera desde la fusión de las tres bolsas bursátiles que

existían en Colombia en el 2001: bolsa de Bogotá, de Medellín y de Occidente. Trimestralmente se recalcula la composición del IGBC y se obtiene una nueva canasta de acciones.

A partir de la información disponible en la página web de la [BVC](#) para las cincuenta y dos empresas inscritas actualmente, se toma como criterio de selección aquellas empresas que presentan una frecuencia de transacción superior al 40% por trimestre. Bajo este criterio se seleccionó un conjunto de dieciocho empresas. La información que se trabajó es el precio de la acción diaria desde el 2 de enero de 2008 hasta el 31 de marzo de 2010, para un total de 7 trimestres. De las dieciocho acciones que en alguno de estos días no fueron transadas, se estimaron estos datos faltantes por medio de la metodología propuesta por [Caporello & Maravall \(2004\)](#) en el programa TSW. Las dieciocho variables consideradas se presentan en el Cuadro 4.3, junto con la cantidad de datos faltantes por cada variable; además, en las Figuras 4.6 y 4.7 se observan las series correspondientes. De este conjunto de acciones se obtienen quince factores comunes de acuerdo con [Peña & Poncela \(2006\)](#). En el Cuadro 4.4 se presentan los resultados de la prueba del número de factores, en los Cuadros A.2, A.3 del Apéndice A, se presentan los valores de la matriz de pesos  $\hat{\mathbf{P}}$  de los factores y en las Figuras 4.8, 4.9 se aprecian las estimaciones de los factores. Allí se observa grandes similitudes entre los factores estimados. Por ejemplo, los factores 2, 4, 5, 8, 9, 10, 14 y 15. Así pues, se aprecian únicamente tres factores, por tanto se modifican los parámetros iniciales en el algoritmo EM y se estiman tres factores, los cuales se aprecian en la Figura 4.10. Claramente el factor B sigue una dinámica similar al IGBC.

Nuevamente, para cada uno de los tres factores y para el IGBC se calculan los puntos de giro y finalmente se calculan las distancias entre los puntos de giro del IGBC a cada uno de los puntos de quiebre de los factores (ver Cuadro 4.5). De las tres comparaciones se encuentra que para el factor B las diferencias son inferiores a 20 días, no obstante se construye el diagrama de perfil coincidente para los tres factores. En la Figura 4.11 se presentan los diagramas de perfil coincidente; en estos diagramas se confirma que el factor B puede ser considerado un índice coincidente para los movimientos de las acciones en la BVC.

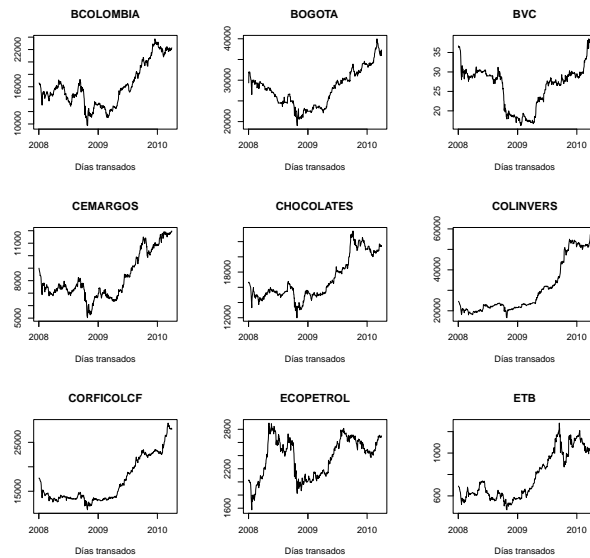


Figura 4.6: Acciones de la BVC seleccionadas en la aplicación

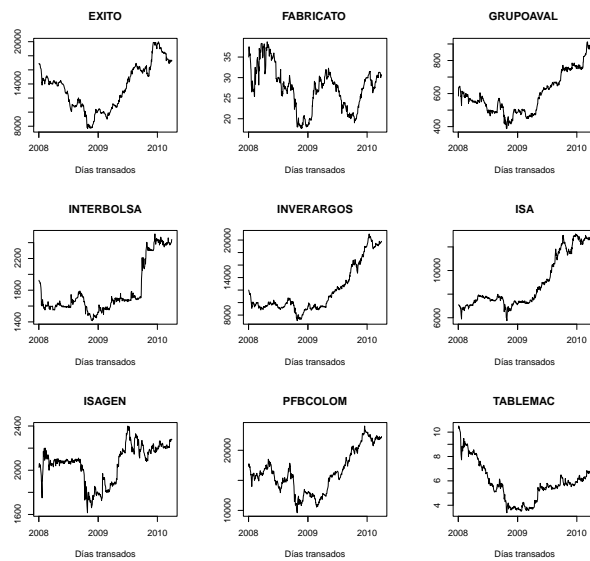


Figura 4.7: Continuación... Acciones de la BVC seleccionadas en la aplicación

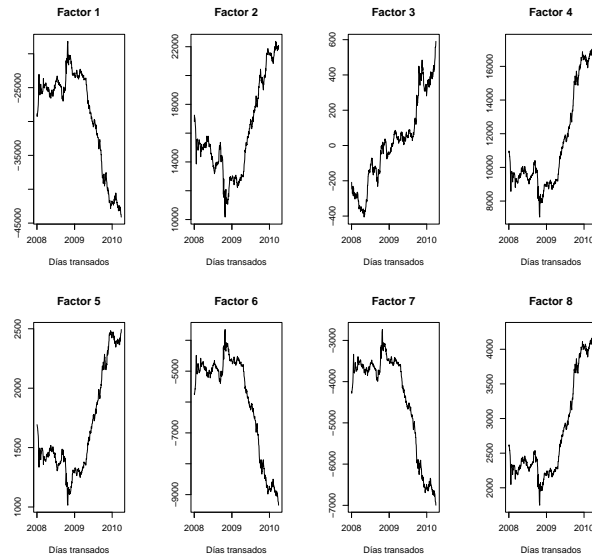


Figura 4.8: Factores estimados con las acciones seleccionados de la BVC

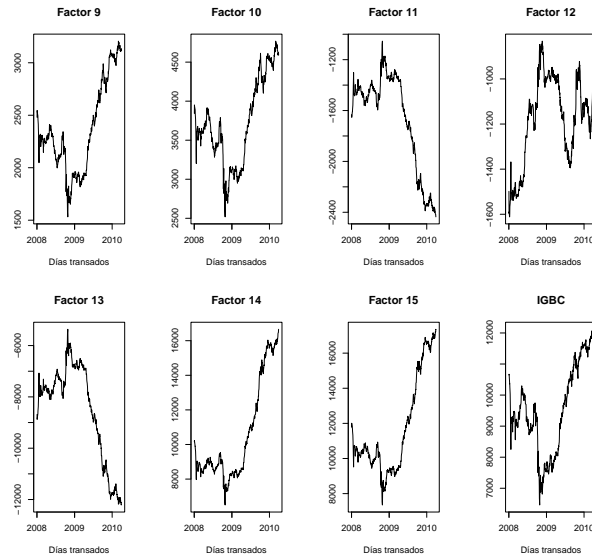


Figura 4.9: Continuación... Factores estimados con las acciones seleccionados de la BVC e IGBC

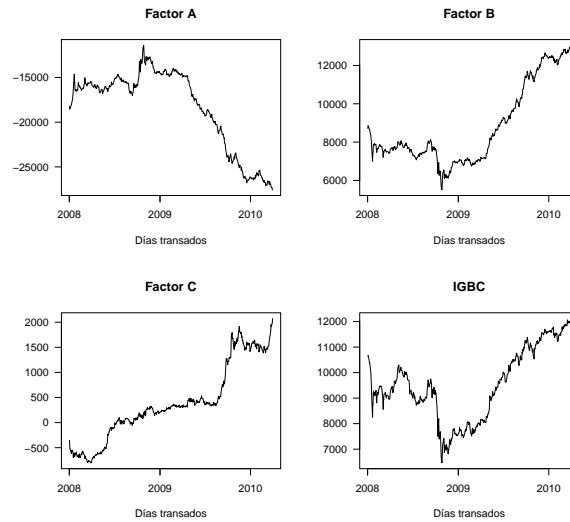


Figura 4.10: Factores estimados definitivos con las acciones seleccionados de la BVC e IGBC

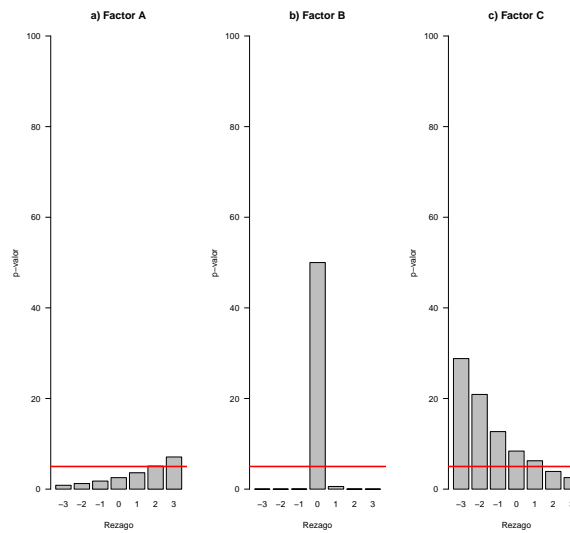


Figura 4.11: Perfil coincidente aplicación acciones de la BVC.

Cuadro 4.3: Acciones consideradas en la aplicación

Nombre Nemotécnico	Cantidad de datos	Faltantes	Total
BCOLOMBIA	545	2	547
BOGOTA	513	34	547
BVC	520	27	547
CEMARGOS	547	0	547
CHOCOLATES	532	15	547
COLINVERS	534	13	547
CORFICOLCF	538	9	547
ECOPETROL	547	0	547
ETB	542	5	547
EXITO	538	9	547
FABRICATO	545	2	547
GRUPOAVAL	540	7	547
INTERBOLSA	540	7	547
INVERARGOS	546	1	547
ISA	547	0	547
ISAGEN	543	4	547
PFBCOLOM	547	0	547
TABLEMAC	540	7	547

El IGBC es un índice que se contruye mediante el promedio ponderado del precio de las acciones que cumplen ciertas condiciones establecidas por la BVC. Dado que, la metodología propuesta hace uso del modelo de factores comunes dinámicos, propuesto por [Peña & Poncela \(2006\)](#)), que parte de la hipótesis de que las variables observables son generadas por medio de un conjunto de factores, siendo estas variables latentes, una de las ventajas de proceder con la metodología propuesta es que los factores y por ende los índices, se calculan teniendo en cuenta la información no solamente en un tiempo  $t$  específico sino también la información anterior a éste punto, caso contrario ocurre con un promedio ponderado donde únicamente se tiene en cuenta la información en el tiempo  $t$ .

Cuadro 4.4: Salida de la prueba (2.16) en la sección 2.3, para los datos de la BVC seleccionados

r	rezago k				
	1	2	3	4	5
0	26531.92 *	22138.49 *	19596.52 *	17982.89 *	16862.05 *
1	21365.53 *	17319.88 *	15005.53 *	13513.48 *	12466.16 *
2	17835.25 *	14130.23 *	11989.95 *	10608.67 *	9612.04 *
3	15203.6 *	11768.09 *	9795.43 *	8505.56 *	7585.79 *
4	12995.27 *	9932.71 *	8146.09 *	6977.40 *	6147.10 *
5	10878.82 *	8143.33 *	6575.35 *	5541.04 *	4828.04 *
6	9073.2 *	6675.09 *	5284.13 *	4366.82 *	3736.75 *
7	7385.93 *	5342.24 *	4145.17 *	3369.27 *	2837.02 *
8	5889.19 *	4143.40 *	3154.99 *	2520.60 *	2079.73 *
9	4608.62 *	3209.99 *	2415.47 *	1902.43 *	1539.06 *
10	3525.29 *	2442.05 *	1792.42 *	1389.85 *	1116.51 *
11	2529.45 *	1708.75 *	1254.12 *	958.09 *	759.98 *
12	1633.12 *	1087.27 *	780.57 *	602.26 *	472.84 *
13	936.44 *	589.62 *	385.89 *	277.73 *	201.70 *
14	316.60 *	213.65 *	152.77 *	120.68 *	86.05 *
15	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00

\* Indica que la hipótesis nula de máximo  $r$  factores es rechazada con un nivel de significancia del 5%.

### 4.3. Sector climatológico

El cambio climático representa una gama de oportunidades y riesgos para el desarrollo humano y por ello se hace necesario estudiarlo. Algunas características como la velocidad y la magnitud de estos cambios, son factores importantes que se deben tener en cuenta ya que muchas de estas modificaciones superan la capacidad de adaptación de las especies. La determinación de los factores responsables del cambio climático puede traer beneficios para la explotación y aprovechamiento de los recursos naturales asegurando con ello la producción de alimentos y desarrollo económico.

Los principales factores responsables del cambio climático se clasifican en tres grupos de acuerdo con su importancia: cambios en la constante solar (ondas



### 4.3 Sector climatológico

Cuadro 4.5: Diferencias en los puntos de giro del IGBC a cada factor

Pto. giro	df1	df2	df3	df4	df5	df6	df7	df8	df9	df10	df11	df12	df13	df14	df15
min		0	-10	-13	-13			-13	0	0		7		-13	-13
max	23	13	-4	2	2	23	23	2	13	13	23	23	23	2	2
min	20	-5	-44	-19	-19	20	20	-19	-5	-5	20	16	20	-19	-19
max	32	-1	-42	-1	-1	32	32	-13	7	6	46	46	32	-13	-1
min	41	0	-10	1	0	41	41	1	0	0	41	80	41	1	1
max	41	1	18	1	8	41	41	1	1	1	41	94	41	1	1
min	31	0	31			31	31	0	0	0	31	119	31	-12	0
max	5	0	-4			5	5	0	0	0	5	73	5	-40	0
min	12	0	12	5	5	-28	-28	0	0	0	12	75	12	-39	0
max	15	-13	15	-13	-13	-24	-24	-13	-13	-2	15	80	15	-36	-13
min	11	0	27	0	0	-12	-12	0	0	0	11	94	11	-23	0
max	12	0	-16	0	0	-11	-11	0	0	0	12	51	12		0
min	11	0	-27	0	0	-83	-84	0	0	0	11	45	11		0
max	83	0	40	-1		-8	-8	0	0	0	83	108	83	0	0
min	8	0	33	0		-66	-66	0	0	0	8	101	8	0	0
max	27	0	43			-49			0	0	27	117	27		
min	-31	0	43			-71			0	0	-23	95	-23		
max	-18	0	58	-8	-8	-80	-18	-8	0	0	-18	110	-18	-8	-8
min	-30	0	54	0	0		-30	0	0	0	-30	51	-30	0	0
max	-13	19	60	19	19		11	19	19	19	-14	40	-13	19	19
min		-1	66	0	0		16	0	-1	-1		38	-18	0	0

electromagnéticas, flujo de partículas cargadas y viento solar), cambios en los parámetros de la órbita terrestre (excentricidad, perihelio afelio y rotación entre otros) y otros factores terrestres (cambios en el nivel de los mares, erupciones volcánicas, vapor de agua, ozono, dióxido de carbono, temperatura de la superficie del mar, aerosoles y procesos en la criósfera).

Aunque los cambios en el clima se le atribuyen a estos tres principales grupos ([Asnani \(1993\)](#)), estos no deben ser considerados por separado a pesar de que las modificaciones que ellos ocasionan sean de magnitudes diferentes.

[Rincón \(2003\)](#) consideró factores pertenecientes al tercer grupo, para los cuales se tiene contacto más directo y cuyos efectos en el clima son percibidos a diario. En dicho trabajo construye un índice del estado del clima, para la Sabana de Bogotá,

por medio de la metodología propuesta por [Stoffer \(1999\)](#) la cual está dirigida a encontrar ciclos que sean compartidos por la mayoría de las series.

Para la aplicación de datos climatológicos y con el fin de construir un índice coincidente del estado del clima, se considera como serie de referencia el índice calculado por [Rincón \(2003\)](#). Para el cálculo de dicho índice se tuvieron en cuenta las siguientes variables: valores totales mensuales de precipitación, número de días mensuales de precipitación, valores máximos de precipitación, valores medios mensuales de temperatura, valores mediamini mensuales de temperatura, valores máximos mensuales de temperatura, valores mínima mensual de temperatura, valores medios mensuales de humedad relativa y valores totales mensuales de evaporación, para un total de 9 variables (ver Figura 4.12). Cada una de estas variables se midió en la Sabana de Bogotá, durante el periodo comprendido entre 1978:01 a 2002:12 (datos mensuales), en la estación denominada Tibaitatá, ubicada en el kilómetro 14 vía municipio de Mosquera.

Al aplicar el procedimiento de [Peña & Poncela \(2006\)](#), para este conjunto de variables, se encontraron cuatro factores comunes. En el Cuadro 4.6 se presentan los resultados de la prueba del número de factores, en el Cuadro A.4 del Apéndice A se presenta la matriz de pesos  $\hat{\mathbf{P}}$  de los factores y en la Figura 4.13 se aprecian las estimaciones de los factores. En estas gráficas se observa que en general tres de los cuatro factores (factores 2, 3 y 4) siguen un comportamiento similar, lo cual indica que al parecer solamente hay dos factores comunes. Dada esta observación se modifican los parámetros iniciales del algoritmo EM de tal manera que sean estimados dos factores, los cuales se presenta en la Figura 4.14 junto con el índice estimado por [Rincón \(2003\)](#).

Luego, se calcularon las distancias entre los puntos de giro de la serie de referencia y los giros de los factores A y B (ver Cuadro 4.7) para construir el diagrama de perfil coincidente (Figura 4.15). En estas gráficas se aprecia que en el rezago cero no se rechaza la hipótesis nula “la suma de las diferencias es cero”, para un nivel de significancia del 5%, para los dos factores; pero el  $p$ -valor en dicho punto del factor B es mayor que el del factor A. Puesto que, para este rezago no se presenta el mayor  $p$ -valor, lo cual indica que, en promedio, los puntos de giro de la serie de referencia se presentan antes que los puntos de giro del factor

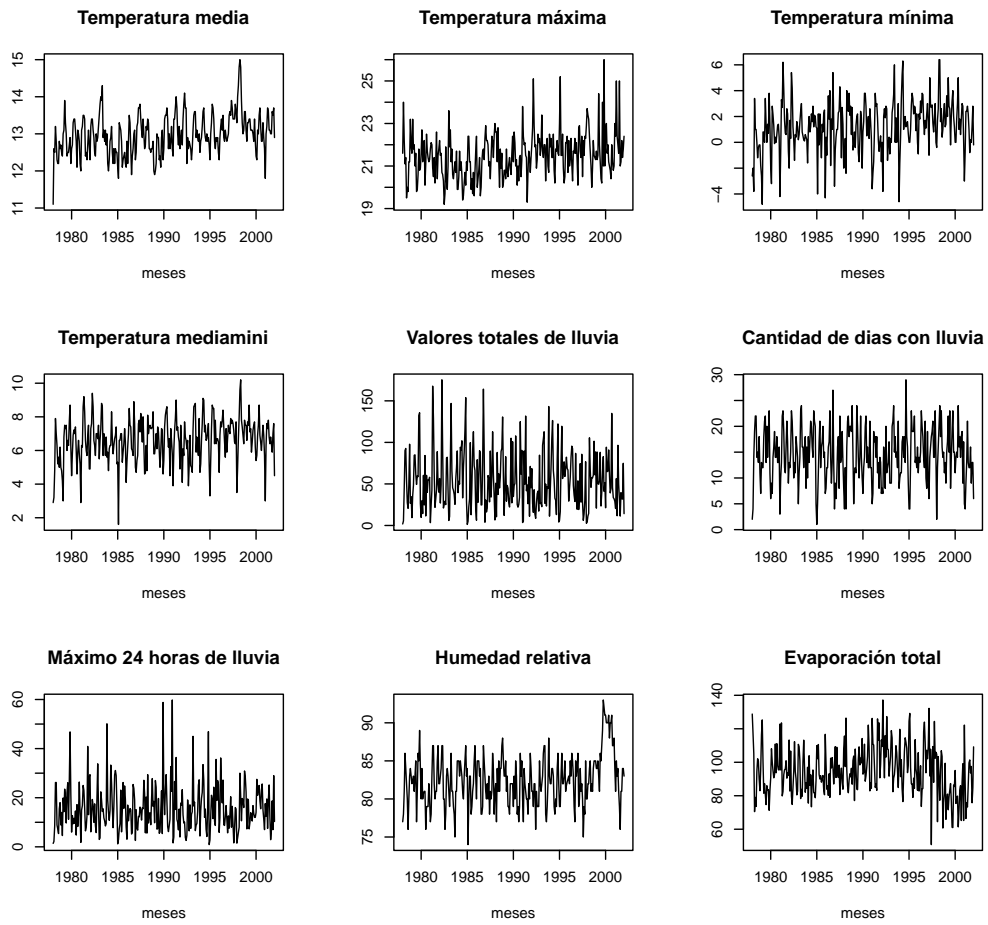


Figura 4.12: Series climatológicas consideradas en la aplicación

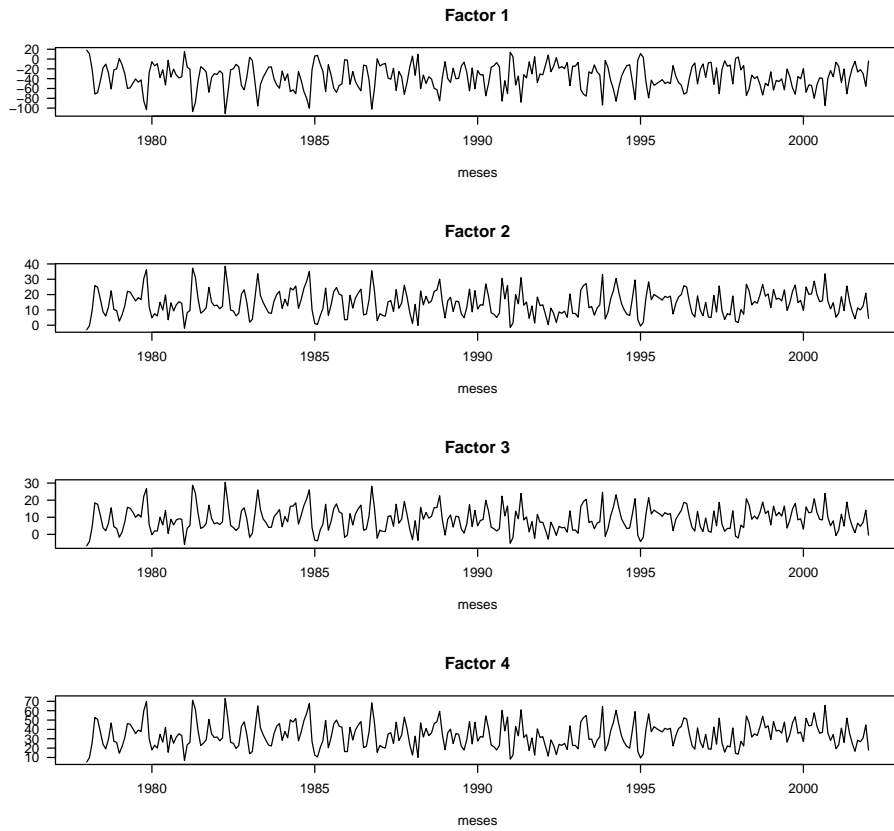


Figura 4.13: Factores estimados con los datos climatológicos en la Sabana de Bogotá

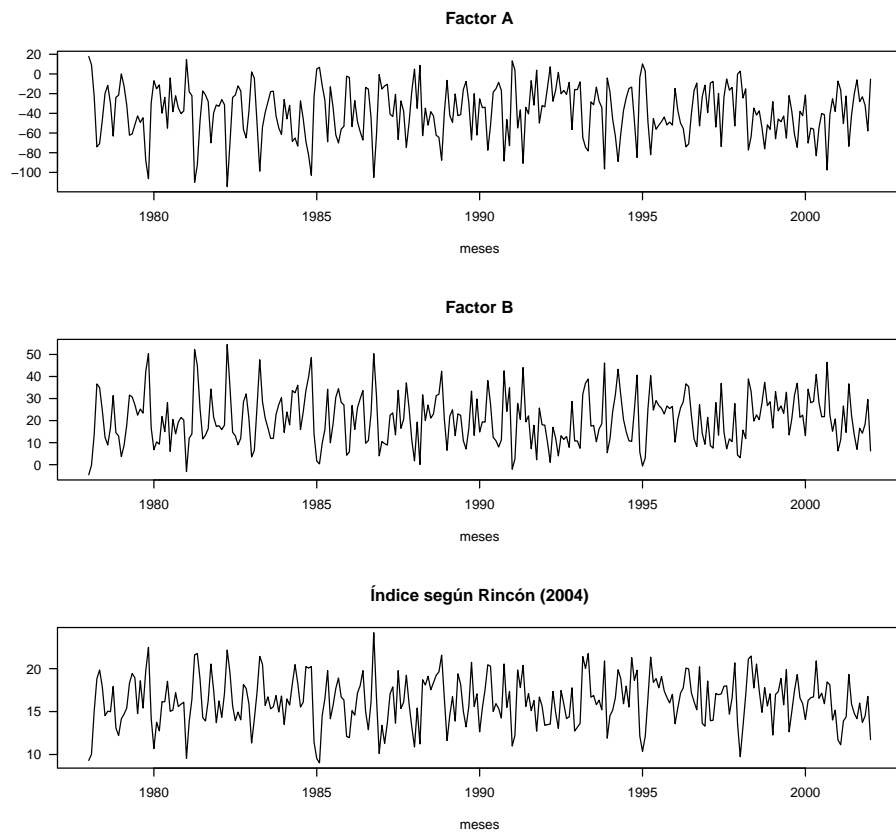


Figura 4.14: Factores comunes definitivos, estimados con los datos climatológicos en la Sabana de Bogotá e índice estimado por [Rincón \(2003\)](#)

Cuadro 4.6: Determinación del número de factores comunes para los datos del clima en la Sabana de Bogotá.

r	rezago k				
	1	2	3	4	5
0	2808.14 *	2568.69 *	2429.72 *	2344.58 *	2304.87 *
1	359.22 *	243.86 *	197.19 *	165.45 *	157.33 *
2	154.45 *	96.77 *	88.09 *	95.02 *	82.63 *
3	63.33 *	61.20 *	52.14 *	52.39 *	34.75
4	29.87	31.77	29.76	28.61	21.85
5	10.78	15.71	11.01	15.01	10.57
6	3.19	9.61	4.40	7.90	3.44
7	0.93	4.49	0.79	1.78	1.45
8	0.00	0.56	0.20	0.09	0.03

\* Indica que la hipótesis nula de máximo  $r$  factores es rechazada con un nivel de significancia del 5%.

B, sin embargo, a la luz de la metodología propuesta se concluye que el índice de [Rincón \(2003\)](#) no es del todo una serie de referencia adecuada.

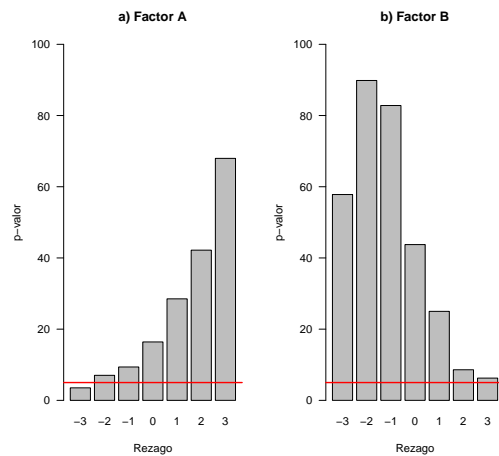


Figura 4.15: Diagramas de Perfil coincidente para el estado del clima

Como en el rezago cero no se rechaza la hipótesis nula, no se descarta del todo la coincidencia entre el factor B y la serie de referencia. Por lo anterior, se

Cuadro 4.7: Diferencias en los puntos de giro del índice del estado del clima según [Rincón \(2003\)](#) al final factor estimado

pto.giro	Fecha pg SR	dcA	dcB
max	nov-79	-	0
min	ene-83	-	-
max	may-84	-9	-
min	feb-85	-7	0
max	nov-88	-	-6
min	dic-92	19	9
max	may-93	14	-6
min	ene-95	14	0
max	may-98	4	-
min	ene-99	3	-
max	may-00	10	-5
min	feb-01	5	-6
max	jun-02	4	-4

Donde pg SR es el punto de giro de la serie de referencia, dcA y dcB las distancias entre los puntos de giro de los Factorres y la serie de referencia.

concluye que el factor B puede considerarse un índice del estado del clima.

Ahora bien, como el índice calculado, en el presente trabajo, y la serie de referencia se obtienen por medio de distintas metodologías, es pertinente aclarar que, en general, los resultados obtenidos son similares. Esto da tranquilidad dado que, a diferencia con las aplicaciones anteriores, en el sector climatológico generalmente no se cuenta con una serie de referencia, como es el caso del IPR en la economía o el IGBC en las finanzas. Sin embargo, se recomienda usar la metodología obtenida en el presente trabajo, ya que ella hace uso de la estimación de factores comunes y no solamente tiene en cuenta la autocorrelación de cada variable sino la correlación entre las series, mientras que la metodología de [Stoffer \(1999\)](#), por ser una metodología en el dominio de las frecuencias, considera las señales comunes entre las variables a distintas frecuencias, basándose en el hecho de considerar la frecuencia con la cual se obtiene la proporción más grande de varianza explicada. Este método es análogo al análisis por componentes principales, el cual solamente considera la varianza explicada pero no la covarianza

que conforma el proceso vectorial. Otra desventaja de la metodología de [Stoffer \(1999\)](#) frente a la de [Peña & Poncela \(2006\)](#) es que solamente se puede utilizar para series de tiempo estacionarias.



# Capítulo 5

## Conclusiones y recomendaciones

La metodología propuesta en este trabajo permite determinar una medida de cercanía entre los factores comunes dinámicos, obtenidos por medio de la metodología de [Peña & Poncela \(2006\)](#), y una serie de referencia. Ella fue probada en varios contextos (económicos, financieros y climatológicos) y en todos estos casos permitió establecer la proximidad entre los puntos de giro de las series en cuestión. Así, fue posible identificar un índice coincidente en cada uno de dichos sectores. Por tanto, la metodología propuesta cumple el objetivo planteado en este trabajo.

Adicionalmente, esta metodología se probó con dos ejercicios simulados cuyos resultados indican que el factor estimado sigue fielmente la dinámica del factor real. Además, se encontró que conforme aumenta la varianza del término de error, del proceso simulado  $\{\mathbf{y}_t\}$ , la distancia entre los puntos de giro del factor estimado respecto a los del factor simulado aumenta, esto se observa en el diagrama de perfil coincidente cuando se aprecia un corrimiento de los  $p$ -valores mayores hacia rezagos diferentes de cero, dificultando con ello la identificación de cercanía entre las series comparadas.

Ahora bien, con base en los resultados obtenidos en este trabajo es posible considerar varias alternativas de desarrollo. Por ejemplo, si se considera que esta metodología funciona siempre que se cuente con una serie de referencia, pero como en algunos casos esto no es posible (como es el caso del sector climatológico), se hace necesario apoyarla con otros tipos de herramientas propias del área de estudio. Por otra parte, teniendo en cuenta que para el cálculo de

---

los puntos de giro se empleó lo propuesto por [Bry & Boschan \(1971\)](#), que se basa en una serie de suavizamientos, es interesante construir los diagramas de perfil coincidente, tomando como referencia los puntos de giro calculados por medio de una metodología menos descriptiva, es decir, que ésta considere pruebas inferenciales para la obtención de dichos puntos. Una alternativa a ello puede ser la propuesta hecha por [Sarlan \(2001\)](#).

# Referencias

- (????). Bolsa de valores de colombia. <http://www.bvc.com.co/>, consulta: Abril 2010. [30](#), [31](#)
- ASNANI, G. (1993). *Tropical Meteorology*. India, Pune University. [37](#)
- BANERJI, A. (1999). The lead profile and others non-parametrics tools to evaluate survey series as leading indicators. *Economic Cycle Research Institute, CIRET Conference Willington New Zealand, March*. [2](#), [8](#), [11](#), [15](#), [16](#)
- BAUTISTA, F. & GÓMEZ, E. (2007). Una exploración de robustez de tres pruebas: dos de permutación y la de Mann-Whitney. *Revista Colombiana de Estadística*, **30**, 177–185. [17](#)
- BRY, G. & BOSCHAN, C. (1971). Cyclical analysis of time series: Selected procedures and computer programs. *National Bureau of Economic Research, New York*, **20**. [II](#), [6](#), [8](#), [9](#), [26](#), [46](#)
- CAPORELLO, G. & MARAVALL, A. (2004). *Manual TSW*. Banco de España. Madrid. [31](#)
- CASTILLO, L. & NIETO, F. (2008). Using the Gibbs sampler for obtaining a coincident economic index: a model-based approach. *Estadística*, **60**, 19–41. [II](#), [1](#), [28](#), [29](#), [30](#)
- DEMPSTER, A., LAIRD, N. & RUBIN, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc.*, **B 39**, 1–38. [14](#)

- LEHMANN, E.L. & ROMANO, J.P. (2005). *Testing Statistical Hypotheses. (Third Edition)*. Springer, Verlag. [8](#)
- LUTKEPOHL, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer. Verlag, New York. [13](#)
- MARTÍNEZ, W. & NIETO, F. (2007). Uso de tendencias comunes en la construcción de un índice coincidente. *Trabajos de grado, departamento de Estadística, Universidad Nacional de Colombia*. [2](#), [23](#), [24](#)
- NIETO, F. (2003). Identifiability of a coincident index model for the Colombian economy. *Borradores Semanales de Economía, Banco de la República*, **242**. [23](#)
- NIETO, F. (2004). A Coincident Index for the State of the Economy. *International Statistical Review, International Statistical Institute*, **72**. [26](#)
- NIETO, F. & MELO, L.F. (2001). Un índice coincidente para la actividad económica colombiana. *Borradores semanales de economía, Banco de la República*, **195**. [1](#), [3](#), [5](#), [23](#)
- PEÑA, P. & PONCELA, P. (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, **136**, 1237–1257. [2](#), [12](#), [13](#), [14](#), [17](#), [20](#), [21](#), [24](#), [28](#), [30](#), [31](#), [35](#), [38](#), [44](#), [45](#)
- RINCÓN, F. (2003). Construcción de un índice de cambio climático para la Sabana de Bogotá. *Trabajo de pregrado Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia*. **III**, [37](#), [38](#), [41](#), [42](#), [43](#)
- SARGENT, T. & SIMS, C. (1977). Business cycle modeling without pretending to have too much a-priori economic theory, in C. Sims et al., New methods in business cycle research. *Minneapolis: Federal Reserve Bank of Minneapolis*. [3](#)
- SARLAN, H. (2001). Cyclical aspects of business cycle turning points. *International Journal of Forecasting*, **17**, 369–382. [46](#)
- STOCK, J. & WATSON, M. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 351–394. [1](#), [3](#)

- STOCK, J. & WATSON, M. (1991). A probability model of the coincident economic indicators. En K. Lahiri and G.H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records*. *New York: Cambridge University Press*, **4**, 63–85. [3](#)
- STOFFER, D. (1999). Detecting Common Signals in Multiple Time Series Using Spectral Envelope. *Journal of the American Statistical Association*, **94**, 1341–1356. [38](#), [43](#), [44](#)
- WELCH, W. & GUTIÉRREZ, L. (1988). Robust Permutation Test for Matched-Pairs Designs. *Journal of the American Statistical Association*, **83** - **402**, 450–455. [17](#)

# Apéndice A

## Apéndice

Cuadro A.1: Estimación de la matriz de pesos  $\mathbf{P}$ , para los datos de la macroeconomía colombiana

	Factor 1	Factor 2
F_P1	0.02	-0.09
F_P6	-0.08	-0.11
PRCEM	0.39	0.16
IPR	0.38	0.14
IEM_OB	-0.36	-0.09
EFECRC	0.38	0.02
ENERGA	0.41	0.13
IMPRES	0.40	0.09
CART_SBR	0.38	0.20

Cuadro A.2: Estimación de la matriz de pesos  $\mathbf{P}$  para la aplicación de finanzas

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
BCOLOMBIA	-0.211	0.250	-0.390	0.232	-0.268	-0.151	0.142	-0.059
BOGOTA	-0.252	0.574	0.358	-0.359	-0.272	0.359	-0.250	0.282
BVC	0.000	0.001	0.000	0.000	-0.001	0.000	0.000	-0.002
CEMARGOS	-0.113	0.080	0.073	0.198	-0.032	-0.186	-0.336	-0.250
CHOCOLATES	-0.173	0.003	-0.015	0.617	0.220	0.482	-0.462	-0.130
COLINVERS	-0.787	-0.497	-0.122	-0.256	0.035	0.186	0.074	0.005
CORFICOLCF	-0.285	0.073	0.675	0.306	-0.096	-0.248	0.424	-0.296
ECOPETROL	-0.010	0.011	-0.027	0.159	-0.025	-0.021	0.065	0.218
ETB	-0.013	0.001	0.012	0.038	0.064	-0.013	0.040	0.089
EXITO	-0.174	0.462	-0.112	-0.171	0.811	-0.040	0.182	-0.142
FABRICATO	0.000	0.001	-0.001	-0.002	-0.001	0.000	-0.001	0.003
GRUPOAVAL	-0.007	0.009	0.009	-0.001	-0.008	0.007	0.002	-0.006
INTERBOLSA	-0.019	-0.002	-0.025	-0.043	-0.037	-0.020	-0.040	-0.065
INVERARGOS	-0.238	0.043	0.001	-0.130	-0.019	-0.652	-0.539	-0.006
ISA	-0.138	-0.020	0.053	0.362	0.150	-0.237	0.149	0.811
ISAGEN	-0.007	0.024	0.003	0.028	0.014	0.025	0.001	0.082
PFBCOLOM	-0.205	0.364	-0.476	0.179	-0.327	-0.005	0.231	-0.074
TABLEMAC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.001

Cuadro A.3: Continuación... Estimación de la matriz de pesos  $\mathbf{P}$  para la aplicación de finanzas

	Factor 9	Factor 10	Factor 11	Factor 12	Factor 13	Factor 14	Factor 15
BCOLOMBIA	-0.059	0.088	0.727	0.125	-0.075	0.039	-0.012
BOGOTA	0.014	-0.009	0.068	0.033	0.021	0.023	-0.014
BVC	0.003	0.007	-0.007	0.001	-0.010	-0.010	-0.020
CEMARGOS	0.818	-0.117	0.019	-0.122	0.017	0.070	0.009
CHOCOLATES	-0.283	0.002	-0.032	0.041	0.003	-0.014	-0.005
COLINVERS	0.098	0.018	0.002	-0.019	0.014	-0.006	-0.002
CORFICOLCF	-0.117	0.039	-0.049	0.006	-0.019	-0.019	-0.012
ECOPETROL	0.244	0.867	-0.176	0.227	0.175	-0.039	-0.004
ETB	-0.086	0.094	0.085	-0.336	0.335	0.901	0.027
EXITO	0.051	0.055	0.010	0.039	-0.013	-0.033	0.001
FABRICATO	-0.003	0.009	-0.005	0.005	0.002	0.019	-0.011
GRUPOAVAL	-0.006	0.007	0.022	-0.020	0.027	-0.014	0.998
INTERBOLSA	0.000	0.002	-0.222	0.556	-0.649	0.372	0.040
INVERARGOS	-0.388	0.097	-0.110	-0.019	0.049	-0.041	-0.002
ISA	0.061	-0.310	-0.016	0.064	-0.084	-0.040	0.010
ISAGEN	-0.002	0.293	0.051	-0.683	-0.642	-0.186	0.005
PFBCOLOM	-0.025	-0.148	-0.600	-0.142	0.073	-0.024	0.004
TABLEMAC	0.000	0.000	-0.002	-0.001	-0.001	-0.001	0.001



---

Cuadro A.4: Estimación de la matriz de pesos  $\mathbf{P}$  para los datos de la aplicación climatológica.

	Factor 1	Factor 2	Factor 3	Factor 4
temp_media	-0.188	0.067	0.301	0.200
temp_maxabs	0.213	-0.043	0.012	0.035
temp_minabs	-0.460	0.140	0.233	0.234
temp_mediaMin	-0.523	0.160	0.289	0.273
hum_med	-0.377	0.154	0.002	0.336
eva_total	0.367	-0.120	-0.088	-0.193
lluv_total	-0.487	0.155	0.136	0.252
n_diasLluv	-0.462	0.135	0.173	0.190
max_lluv24h	-0.318	0.107	0.062	0.185