

Una metodología multivariada de
desagregación temporal

Por: Jorge Luis Hurtado Guarín
Luis Fernando Melo Velandia

Borradores de ECONOMÍA

Núm. 586

2010



tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col

UNA METODOLOGÍA MULTIVARIADA DE DESAGREGACIÓN TEMPORAL *

JORGE L. HURTADO
LUIS F. MELO

BANCO DE LA REPÚBLICA

RESUMEN. En este artículo se propone una extensión de la metodología multivariada de desagregación temporal de Di-Fonzo [1990]. Esta supone que los errores de las series de alta frecuencia siguen un modelo VAR(1) en lugar de un proceso ruido blanco. Adicionalmente, se realiza una reseña de diferentes métodos de desagregación, tanto univariados como multivariados, incluyendo sus principales ventajas y desventajas. Finalmente, se lleva a cabo una aplicación multivariada para obtener las cuentas nacionales colombianas mensuales a partir de datos trimestrales.

Palabras claves. Desagregación temporal, restricciones de agregación temporales y contemporáneas.

Códigos de clasificación JEL. C32, C51, E01.

1. INTRODUCCIÓN

La mayor cantidad de información con la que actualmente se cuenta, permite que se tenga una mejor disponibilidad de datos de alta frecuencia. Sin embargo, todavía existen algunos casos en los que no se cuenta con observaciones de algunas variables a la frecuencia deseada. Un ejemplo típico es el PIB colombiano, que se encuentra disponible en frecuencias trimestrales o anuales mientras que algunos estudios requieren que sus valores sean mensuales.

En este punto la desagregación temporal adquiere relevancia puesto que permite estimar valores de alta frecuencia dadas las observaciones de baja frecuencia. En la literatura se encuentran diferentes métodos de desagregación temporal univariados y multivariados. Algunas de estas metodologías utilizan variables adicionales a la variable de interés para estimar la serie requerida de alta frecuencia. Estas variables son incluidas para tener información de la dinámica de alta frecuencia de la serie a desagregar.

Fecha: Febrero, 2010.

* Los resultados y opiniones son responsabilidad exclusiva de los autores y su contenido no compromete al Banco de la República ni a su junta directiva. Los autores agradecen los valiosos comentarios y sugerencias de Eliana González y Daniel Quintero, así como a la sección de cuentas financieras de la SG-EE del Banco de la República por la base de datos proporcionada.

A medida que se cuenta con más información, las estimaciones de las series desagregadas pueden mejorarse. Desde este punto de vista, los métodos multivariados tienen una ventaja sobre los univariados, ya que incluyen un mayor número de datos y relaciones en las metodologías de estimación. En este artículo se realiza una extensión de la metodología multivariada de Di-Fonzo [1990], a la vez que se describen algunos métodos de desagregación univariados y multivariados. Adicionalmente, se aplican estas metodologías a las cuentas nacionales colombianas.

El artículo está organizado de la siguiente forma, en la segunda sección se define el problema de desagregación temporal y se indican las convenciones de notación que seguirá el artículo. En la tercera sección se describen algunos métodos de desagregación univariados y se discuten sus ventajas y desventajas. En la cuarta sección se introducen algunos métodos multivariados, y se propone una extensión a una de estas metodologías. La aplicación de los métodos multivariados se efectúa en la quinta sección. Por último, en la sexta sección se presentan algunos comentarios finales.

2. EL PROBLEMA DE DESAGREGACIÓN TEMPORAL

La desagregación temporal consiste en generar una serie de alta frecuencia partiendo de la misma en baja frecuencia. Algunos de estos métodos utilizan una o más variables adicionales a la variable de interés, estas variables se denominan *variables indicadoras*.¹ En general, se cuenta con una serie de baja frecuencia observada en un periodo muestral determinado, que se desea desagregar de tal forma que la observación de cada periodo se distribuya en una cantidad fija de subperiodos. Por ejemplo, una serie anual observada durante 10 años se podría desagregar para tener 120 observaciones mensuales o 40 observaciones trimestrales. En el primer caso el número de subperiodos es 12, y en el segundo es 4.

En la literatura sobre este tema se encuentran diversos métodos de desagregación que se pueden clasificar inicialmente como univariados y multivariados. Los métodos univariados son aquellos en los que hay una única serie a desagregar, mientras que en los multivariados hay más de una serie a desagregar simultáneamente. Adicionalmente, dentro de los métodos univariados y multivariados se pueden encontrar los que usan variables indicadoras y los que no.

Existen tres formas de desagregación temporal que dependen de la naturaleza de la variable que se desea desagregar. Éstas son: distribución, interpolación y extrapolación. A continuación se describe cada una de ellas.

Cuando se tiene una variable flujo en baja frecuencia, y se quiere desagregar para obtenerla en alta frecuencia, la suma de los valores de los subperiodos de alta frecuencia de cada periodo debe ser igual al valor del periodo correspondiente de baja frecuencia. En este caso se tiene un problema de *distribución*. Un ejemplo de este problema es el de la desagregación del PIB trimestral a frecuencia mensual, donde se requiere que la suma de los valores mensuales estimados para un trimestre sea igual al valor observado en dicho trimestre. Por otro lado, si la observación trimestral es un

¹Estas se pueden entender como variables *proxy* de alta frecuencia de la serie a desagregar.

promedio de las observaciones mensuales, la estimación de las series mensuales también se puede ver como un problema de distribución.

Un problema de *interpolación* surge cuando se cuenta con una serie de baja frecuencia de una variable stock, observada al inicio o al final del periodo, y se quiere estimar la serie de alta frecuencia de la misma variable. En este caso, si la variable es stock observada al final (inicio) del periodo, se requiere que el valor de la serie estimada de alta frecuencia en el último (primer) subperiodo de cada periodo sea igual a la observada en el periodo correspondiente de la serie de baja frecuencia. Un ejemplo de una variable stock observada al final del periodo, es el acervo anual de capital. En el caso que se desee trimestralizar esta variable, se requiere que el valor del último trimestre de cada año sea igual al valor anual correspondiente.

Después de desagregar una serie por distribución o interpolación, en algunas ocasiones se necesita obtener estimaciones que estén por fuera de la muestra observada. En este caso se tiene un problema de *extrapolación*. Un ejemplo de un problema de extrapolación es el de estimar el PIB mensual desde 2010 : 1 hasta 2010 : 3 a partir del PIB trimestral observado desde 1990 : 1 hasta 2009 : 4.

Después de esta introducción breve al problema de desagregación temporal, es necesario formalizar la discusión con el fin de entender en detalle los diferentes métodos. A continuación se presentan las convenciones de notación utilizadas a lo largo del artículo, cuya estructura es similar a la de Di-Fonzo [2002].

2.1. Notación. Se supone que se ha observado una variable a una frecuencia baja durante S periodos. Sea Y_s la observación de la variable en baja frecuencia en el periodo s , con $s = 1, \dots, S$, y sea \mathbf{Y} un vector de dimensión $(S \times 1)$ donde $\mathbf{Y} = [Y_1, \dots, Y_S]'$. Suponga, además, que cada periodo se puede dividir en un número finito, M , de subperiodos de alta frecuencia.

Cuando se utilizan metodologías de desagregación que requieren el uso de variables indicadoras, se define $\mathbf{w}_{m,s}$, donde $\mathbf{w}_{m,s}$ es un vector $(1 \times p)$ de observaciones de alta frecuencia de $p \geq 1$ variables indicadoras en el periodo s y subperiodo m , y sea \mathbf{W}_s el vector de las observaciones de baja frecuencia del mismo vector de variables indicadoras en el periodo s , que consiste en la agregación de $\mathbf{w}_{m,s}$ en s y los subperiodos $m = 1, \dots, M$. Adicionalmente, se define a W como la matriz de dimensión $(S \times p)$ que contiene las observaciones de baja frecuencia de todos los periodos de las p variables indicadoras. Por otro lado, si se define $\mathbf{w}^h = [w_1^h, \dots, w_N^h]'$ como el vector de observaciones de baja frecuencia de la variable indicadora h , con $h = 1, \dots, p$, entonces $w = [\mathbf{w}^1, \dots, \mathbf{w}^p]$ es la matriz de dimensión $(N \times p)$ que contiene todas las observaciones de alta frecuencia de las p variables indicadoras en N subperiodos, con $N \geq M S$.

A partir de los datos observados \mathbf{Y} y w , se busca estimar los valores $y_{m,s}$ que corresponden a los valores no observados de alta frecuencia en el subperiodo m del periodo s de la variable de interés. Sean $\mathbf{y} = [y_1, \dots, y_N]'$ y $\hat{\mathbf{y}} = [\hat{y}_1, \dots, \hat{y}_N]'$ los vectores que contienen los valores no observados de alta frecuencia y los valores estimados de alta frecuencia de la variable de interés, respectivamente.

Los dos vectores son de dimensión $(N \times 1)$.² En algunos métodos se utiliza un estimador preliminar de \mathbf{y} , que se nota como $\tilde{\mathbf{y}}$ y tiene la misma dimensión que $\hat{\mathbf{y}}$.

Ejemplo 1. Con el objeto de clarificar las definiciones anteriores, suponga que se cuenta con dos observaciones anuales de una variable, Y_1 y Y_2 , que se quieren desagregar a frecuencia trimestral, y además se tienen observaciones trimestrales de dos variables indicadoras, entonces el número de periodos de baja frecuencia (años) es $S = 2$, y el número de subperiodos de alta frecuencia (número de trimestres por año) de cada periodo es $M = 4$. Por lo tanto, el número total de subperiodos de alta frecuencia (número de trimestres en los dos años) es $N = 4 \times 2 = 8$. En este caso, existen ocho vectores fila, $\mathbf{w}_{1,1}, \dots, \mathbf{w}_{4,1}, \mathbf{w}_{1,2}, \dots, \mathbf{w}_{4,2}$, de dimensión (1×2) que contienen las observaciones de las dos variables indicadoras de cada trimestre para cada uno de los dos años. Los vectores columna de observaciones de alta frecuencia de las dos variables indicadoras son: $\mathbf{w}^1 = [w_1^1, \dots, w_8^1]'$ y $\mathbf{w}^2 = [w_1^2, \dots, w_8^2]'$, donde w_n^h es la observación de la variable indicadora h en el trimestre n , donde $h = 1, 2$ y $n = 1, \dots, 8$.

Si se supone que la forma de agregar las observaciones trimestrales de las variables indicadoras es sumándolas, entonces los vectores que contienen las observaciones anuales de las variables indicadoras son $\mathbf{W}_1 = [\sum_{n=1}^4 w_n^1, \sum_{n=1}^4 w_n^2]$ y $\mathbf{W}_2 = [\sum_{n=5}^8 w_n^1, \sum_{n=5}^8 w_n^2]$. Por otro lado, las matrices y vectores \mathbf{Y} , W y w , definidas al comienzo de esta sección están dadas por:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix},$$

$$W = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^4 w_n^1 & \sum_{n=1}^4 w_n^2 \\ \sum_{n=5}^8 w_n^1 & \sum_{n=5}^8 w_n^2 \end{bmatrix},$$

y

$$w = [\mathbf{w}^1 \quad \mathbf{w}^2] = \begin{bmatrix} w_1^1 & w_1^2 \\ w_2^1 & w_2^2 \\ w_3^1 & w_3^2 \\ w_4^1 & w_4^2 \\ w_5^1 & w_5^2 \\ w_6^1 & w_6^2 \\ w_7^1 & w_7^2 \\ w_8^1 & w_8^2 \end{bmatrix}$$

En este caso, W y \mathbf{Y} hacen parte de la notación de las series agregadas (de frecuencia baja), mientras que w está asociada a los datos no agregados (de frecuencia alta).

²En general $N = M S$, pero en el caso de extrapolación $N > M S$.

Como se explicó al inicio de la sección, dependiendo de la naturaleza de la variable las condiciones del problema de desagregación cambian. En particular, la *restricción de agregación* es distinta si la variable a desagregar es un flujo, un stock o un promedio. Sea \mathbf{c} un vector de dimensión $M \times 1$, de tal forma que al premultiplicar los valores de los subperiodos, de un periodo determinado, por dicho vector se obtiene la observación de baja frecuencia correspondiente. Entonces se define la restricción de agregación, para el periodo s , como:

$$Y_s = \mathbf{c}' [y_{1,s} \quad \dots \quad y_{M,s}]' \quad (1)$$

donde $\mathbf{c}' = [c_1, \dots, c_M]$ es un vector de M constantes conocidas. Si Y_s representa un flujo, $\mathbf{c}' = [1, 1, \dots, 1]$; si Y_s es un promedio, $\mathbf{c}' = [1/M, 1/M, \dots, 1/M]$; si es un stock observado al final del periodo, $\mathbf{c}' = [0, 0, \dots, 1]$, y si es un stock observado al inicio del periodo, $\mathbf{c}' = [1, 0, \dots, 0]$. Los primeros dos casos corresponden al problema de distribución, y los dos últimos corresponden al de interpolación. Por lo tanto, la restricción de agregación para el periodo s puede tomar las siguientes formas:

Distribución:

- Flujo: $Y_s = [1 \quad 1 \quad \dots \quad 1][y_{1,s} \quad \dots \quad y_{M,s}]' = \sum_{m=1}^M y_{m,s}$
- Promedio: $Y_s = \frac{1}{M}[1 \quad 1 \quad \dots \quad 1][y_{1,s} \quad \dots \quad y_{M,s}]' = \frac{1}{M} \sum_{m=1}^M y_{m,s}$

Interpolación:

- Stock (Inicio del Periodo): $Y_s = [1 \quad 0 \quad \dots \quad 0][y_{1,s} \quad \dots \quad y_{M,s}]' = y_{1,s}$
- Stock (Final del Periodo): $Y_s = [0 \quad 0 \quad \dots \quad 1][y_{1,s} \quad \dots \quad y_{M,s}]' = y_{M,s}$

Es conveniente definir el conjunto de restricciones de agregación de todos los periodos, con el fin de asociar la restricción de agregación a los vectores \mathbf{Y} y \mathbf{y} . Sea C_1 la matriz de dimensión $S \times MS$ definida como:

$$C_1 = I_S \otimes \mathbf{c}' = \begin{bmatrix} c_1 & \dots & c_M & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & c_1 & \dots & c_M & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & c_1 & \dots & c_M \end{bmatrix} \quad (2)$$

donde I_S es una matriz identidad de orden S . Para los problemas de distribución e interpolación, sin extrapolación, el conjunto de restricciones de agregación se define como $\mathbf{Y} = C_1 \mathbf{y}$. En el caso de extrapolación a un horizonte r , el conjunto de restricciones de agregación es:

$$\mathbf{Y} = C_2 \mathbf{y} \quad (3)$$

donde $C_2 = C_1 | 0_r$ es la matriz formada por la unión (por columnas) de las matrices C_1 y 0_r , siendo esta última una matriz de ceros de dimensión $M \times r$.³ La matriz de ceros se incluye porque cuando

³Esta matriz ceros implica que no hay restricciones de agregación en el periodo a extrapolar.

se asocia un problema de extrapolación, los valores de alta frecuencia estimados por fuera de la muestra no están sujetos a la restricción de agregación.

Para simplificar la notación, se define C como la matriz de agregación, que será igual a C_1 si se tiene un problema de distribución o interpolación sin extrapolación, y será igual a C_2 si se tiene un problema de extrapolación. Dado lo anterior, el conjunto de restricciones de agregación es:

$$\mathbf{Y} = C\mathbf{y} \quad (4)$$

La matriz de agregación también se puede utilizar para agregar otras variables de alta frecuencia. Por ejemplo, se puede escribir el conjunto de todas las observaciones de baja frecuencia de las variables indicadoras como $W = Cw$.

Ejemplo 2. Suponga que el problema de desagregación consiste en estimar la serie \mathbf{y} asociada a la serie anual \mathbf{Y} dada en el Ejemplo 1, a través de la solución de un problema de distribución sin extrapolación. Es decir, la suma de los cuatro datos trimestrales de la serie desagregada es igual al dato anual correspondiente y, adicionalmente, no se incluyen observaciones por fuera de la muestra observada. De esta forma, los vectores y matrices \mathbf{y} , $\hat{\mathbf{y}}$, \mathbf{c} y C definidos anteriormente son los siguientes:

$$\mathbf{y} = \begin{bmatrix} y_{1,1} \\ y_{2,1} \\ y_{3,1} \\ y_{4,1} \\ y_{1,2} \\ y_{2,2} \\ y_{3,2} \\ y_{4,2} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix}, \quad \hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_{1,1} \\ \hat{y}_{2,1} \\ \hat{y}_{3,1} \\ \hat{y}_{4,1} \\ \hat{y}_{1,2} \\ \hat{y}_{2,2} \\ \hat{y}_{3,2} \\ \hat{y}_{4,2} \end{bmatrix} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \\ \hat{y}_5 \\ \hat{y}_6 \\ \hat{y}_7 \\ \hat{y}_8 \end{bmatrix},$$

$$\mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Teniendo en cuenta las anteriores convenciones de notación, en la siguiente sección se explican algunos métodos de desagregación.

3. MÉTODOS UNIVARIADOS DE DESAGREGACIÓN TEMPORAL

En esta sección se hace una reseña de algunos de los métodos univariados de desagregación temporal más utilizados en la literatura. Adicionalmente, se discuten sus ventajas y desventajas. En la Tabla 1 se resumen las características principales de estos métodos.

3.1. Método de Denton [1971]. El objetivo del artículo de Denton [1971] es encontrar la forma de ajustar series preliminares (observadas) mensuales o trimestrales para que coincidan con

Método	Relación de alta frecuencia	Supuestos sobre el error	Desventajas
Denton [1971]	-	-	<ul style="list-style-type: none"> No tiene en cuenta las características estocásticas de las series preliminar y de alta frecuencia.
Chow y Lin [1971]	$y_t = \mathbf{w}_t' \boldsymbol{\beta} + u_t$	$u_t \sim \text{AR}(1)$	<ul style="list-style-type: none"> El supuesto sobre el error es muy simplista.
Fernández [1981]	$y_t = \mathbf{w}_t' \boldsymbol{\beta} + u_t$	$u_t = u_{t-1} + \varepsilon_t$, con ε_t ruido blanco	<ul style="list-style-type: none"> No permite que exista una relación de cointegración entre y_t y \mathbf{w}_t.
Litterman [1983]	$y_t = \mathbf{w}_t' \boldsymbol{\beta} + u_t$	$u_t \sim \text{ARIMA}(1, 1, 0)$	<ul style="list-style-type: none"> No soluciona el problema de Fernández [1981] asociado a la cointegración entre y_t y \mathbf{w}_t. No desarrolla un método general para estimar el parámetro autorregresivo.
Guerrero [1990]	$Y_s = \mathbf{W}_s' \boldsymbol{\beta} + U_s$ $y_t = \mathbf{w}_t' \boldsymbol{\beta}$ $\tilde{y}_t \sim \text{ARIMA}(p, d, q)$ $y_t \sim \text{ARIMA}(p, d, q)$	-	<ul style="list-style-type: none"> El método no se puede aplicar cuando existe cointegración entre y_t y \mathbf{w}_t porque se violan algunos de los supuestos de la metodología.
Santos Silva y Cardoso [2001]	$y_t = ky_{t-1} + \mathbf{w}_t' \boldsymbol{\beta} + u_t$	u_t es un ruido blanco	<ul style="list-style-type: none"> Sólo incluye un rezago de la variable dependiente en la relación de alta frecuencia, y no considera modelos más generales.

En estos métodos y_t corresponde a los valores de la serie de alta frecuencia que se desea estimar, $t = 1, \dots, T$; Y_s son los valores observados de baja frecuencia, con $s = 1, \dots, S$; \mathbf{w}_t y \mathbf{W}_s son los valores de la serie indicadora en alta y baja frecuencia, respectivamente; \tilde{y}_t son los valores de una estimación preliminar de y_t ; u_t y ε_t son errores de alta frecuencia cuya especificación difiere según cada método. Finalmente, U_s corresponde al error de baja frecuencia.

TABLA 1. Principales características de algunos métodos univariados de desagregación temporal.

agregados o promedios anuales de fuentes independientes. El método de Denton no permite extrapolación, ya que exige que $N = MS$.

El método busca estimar una serie de alta frecuencia ($\hat{\mathbf{y}}$) a partir de una serie preliminar observada ($\tilde{\mathbf{y}}$), de la misma frecuencia, utilizando restricciones de agregación dadas por la serie observada de baja frecuencia (\mathbf{Y}).

El problema de Denton consiste en minimizar una función de pérdida, $\rho(\mathbf{y}, \tilde{\mathbf{y}})$, que se define como una forma cuadrática sobre la diferencia entre la serie de alta frecuencia a estimar y la serie preliminar, sujeta a la restricción de agregación:

$$\min_{\mathbf{y}} \rho(\mathbf{y}, \tilde{\mathbf{y}}) = (\mathbf{y} - \tilde{\mathbf{y}})' A (\mathbf{y} - \tilde{\mathbf{y}}) \quad \text{s.a.} \quad \mathbf{C}\mathbf{y} = \mathbf{Y} \quad (5)$$

donde A es una matriz simétrica de dimensión $MS \times MS$ cuya forma se especificará más adelante. El lagrangiano asociado a la expresión (5) es:

$$\mathcal{L} = (\mathbf{y} - \tilde{\mathbf{y}})'A(\mathbf{y} - \tilde{\mathbf{y}}) - 2\boldsymbol{\lambda}'(\mathbf{Y} - C\mathbf{y}) \quad (6)$$

donde $\boldsymbol{\lambda}' = [\lambda_1, \lambda_2, \dots, \lambda_S]$. Derivando (6) con respecto a \mathbf{y} y $\boldsymbol{\lambda}$, e igualando a cero, se obtiene:

$$\begin{bmatrix} \hat{\mathbf{y}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} A & C' \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ C & I_S \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{y}} \\ \mathbf{Y} - C\tilde{\mathbf{y}} \end{bmatrix} \quad (7)$$

donde I_S es la matriz identidad de orden S . Resolviendo para $\hat{\mathbf{y}}$, se tiene el siguiente resultado:

$$\hat{\mathbf{y}} = \tilde{\mathbf{y}} + K(\mathbf{Y} - C\tilde{\mathbf{y}}) \quad (8)$$

$$K = A^{-1}C'(CA^{-1}C')^{-1} \quad (9)$$

La ecuación (8) muestra que la estimación de \mathbf{y} es una combinación lineal entre la serie preliminar $\tilde{\mathbf{y}}$ y la discrepancia entre las agregaciones de la serie preliminar $C\tilde{\mathbf{y}}$ y la serie agregada \mathbf{Y} . Es decir, la estimación preliminar se ajusta distribuyendo la discrepancia agregada de acuerdo a la matriz de ponderaciones K . Sin embargo, K depende de A , por lo que se debe tener una forma explícita de A con el fin de realizar la estimación.

Una primera alternativa es definir $A = I_{MS}$, en este caso la función de pérdida es la suma de los cuadrados de las diferencias entre la serie preliminar y la serie a estimar, lo que implica que la discrepancia se está distribuyendo en montos iguales entre los M subperiodos.

Denton propone una especificación alterna de A , de tal forma que la función de pérdida es la suma de los cuadrados de las d -ésimas diferencias de $\mathbf{y} - \tilde{\mathbf{y}}$, es decir, $\rho(\mathbf{y}, \tilde{\mathbf{y}}) = \sum_{i=1}^{MS} \Delta^d(y_i - \tilde{y}_i)^2$. Para obtener esta función, es necesario definir primero una matriz D de dimensión $MS \times MS$ de la forma:

$$D = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix} \quad (10)$$

Entonces, el vector de primeras diferencias de $\mathbf{y} - \tilde{\mathbf{y}}$ se puede escribir como $D(\mathbf{y} - \tilde{\mathbf{y}})$. Si se define $A = D'D$, entonces la suma de los cuadrados de las primeras diferencias de $\mathbf{y} - \tilde{\mathbf{y}}$ está dada por:

$$\rho(\mathbf{y}, \tilde{\mathbf{y}}) = \sum_{i=1}^{MS} \Delta(y_i - \tilde{y}_i)^2 = (\mathbf{y} - \tilde{\mathbf{y}})'D'D(\mathbf{y} - \tilde{\mathbf{y}}).$$

Si se quiere un orden de diferenciación mayor, se debe premultiplicar el vector de primeras diferencias por otra matriz D . Así, para obtener la d -ésima diferencia de $\mathbf{y} - \tilde{\mathbf{y}}$, se premultiplica el vector por el producto de d matrices D : $D^d(\mathbf{y} - \tilde{\mathbf{y}})$. En este último caso, la matriz A tiene la forma $(D')^d D^d$, y la función de pérdida:

$$\rho(\mathbf{y}, \tilde{\mathbf{y}}) = \sum_{i=1}^{MS} \Delta^d (y_i - \tilde{y}_i)^2 = (\mathbf{y} - \tilde{\mathbf{y}})' (D')^d D^d (\mathbf{y} - \tilde{\mathbf{y}}) \quad (11)$$

Como señala Denton, la matriz D simplifica el cálculo de la inversa de A . Sin embargo, el método tiene algunas debilidades. La principal es que la función de pérdida se minimiza en cero, lo que implica que $\rho(\mathbf{y}, \tilde{\mathbf{y}})$ es mínima cuando $\mathbf{y} = \tilde{\mathbf{y}}$. Dicha condición es una restricción importante, porque implica que en el punto óptimo la estimación preliminar y la estimación final deben ser iguales. Por otro lado, el método no tiene en cuenta las características estocásticas de las series preliminar y final. Por ejemplo, el orden de integración o si existe una relación de cointegración entre estas series.

3.2. Método de Chow y Lin [1971]. A diferencia del método de Denton [1971], Chow y Lin [1971] resuelven el problema de desagregación temporal encontrando el mejor estimador lineal insesgado (MELI) de la serie a estimar \mathbf{y} . Una ventaja de este método con respecto al de Denton [1971], es que éste si permite extrapolar.

Para obtener el mejor estimador lineal insesgado de \mathbf{y} , Chow y Lin suponen que existe una relación lineal entre la serie a estimar (\mathbf{y}) y las p series indicadoras (w) en alta frecuencia:

$$\mathbf{y} = w\boldsymbol{\beta} + \mathbf{u} \quad (12)$$

donde \mathbf{y} y w son el vector y la matriz definidas en la sección anterior, $\boldsymbol{\beta}$ es un vector de coeficientes de dimensión $p \times 1$, y \mathbf{u} es un proceso *i.i.d.* de dimensión $N \times 1$ con media cero y matriz de covarianzas V . Si se premultiplica la ecuación (12) por la matriz de agregación C , se obtiene el modelo de baja frecuencia:

$$\mathbf{Y} = C\mathbf{y} = Cw\boldsymbol{\beta} + C\mathbf{u} = W\boldsymbol{\beta} + \boldsymbol{\mu} \quad (13)$$

donde $\boldsymbol{\mu} = C\mathbf{u}$ con $E(\boldsymbol{\mu}\boldsymbol{\mu}') = CVC'$. Chow y Lin muestran que el mejor estimador lineal (MELI) de \mathbf{y} está dado por:

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + VC'(CVC')^{-1}(\mathbf{Y} - W\hat{\boldsymbol{\beta}}) \quad (14)$$

donde $\hat{\boldsymbol{\beta}}$ es el estimador de mínimos cuadrados generalizados del modelo (13), el cual tiene la siguiente forma:

$$\hat{\boldsymbol{\beta}} = [W'(CVC')^{-1}W]^{-1}W'(CVC')^{-1}\mathbf{Y} \quad (15)$$

La ecuación (14) muestra que la solución de Chow y Lin es una combinación lineal entre las series indicadoras w , y el residuo de la ecuación (13). Además, la estimación depende de la forma de la matriz de covarianzas de los errores de alta frecuencia V , que no es observable y se debe estimar suponiendo alguna estructura en los errores de alta frecuencia \mathbf{u} .

En el caso más sencillo, se supone que los errores de alta frecuencia son ruido blanco con varianza σ^2 , lo que implica que $V = \sigma^2 I_N$. Con esta especificación, la discrepancia agregada de cada periodo (residuales de la ecuación (14)) se asigna en partes iguales para cada subperiodo en el caso de distribución. En el caso de interpolación con variable observada al inicio (final) del periodo,

la discrepancia agregada de cada periodo se asigna al primer (último) subperiodo de dicho periodo.

Dadas las limitaciones de la especificación de ruido blanco en los errores de alta frecuencia, Chow y Lin suponen, de forma alterna, que los errores siguen un proceso autorregresivo de orden uno, AR(1), en el cual:

$$u_t = \alpha u_{t-1} + \varepsilon_t \quad E(\varepsilon_t \varepsilon_s) = \delta_{ts} \sigma^2 \quad (16)$$

donde ε_t es un proceso ruido blanco con varianza σ^2 . Adicionalmente, $|\alpha| < 1$, por consiguiente la matriz de covarianzas está definida de la siguiente forma:

$$V = E(\mathbf{u}\mathbf{u}') = \frac{\sigma^2}{1 - \alpha^2} \begin{bmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{N-1} \\ \alpha & 1 & \alpha & \dots & \alpha^{N-2} \\ \alpha^2 & \alpha & 1 & \dots & \alpha^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{N-1} & \alpha^{N-2} & \alpha^{N-3} & \dots & 1 \end{bmatrix} \quad (17)$$

Debido a que V está en función de α y σ^2 , la estimación de la matriz de covarianzas implica la estimación de estos parámetros. Chow y Lin proponen un método iterativo para estimar α , sin embargo, estos autores sólo lo desarrollan para el caso de la desagregación de series trimestrales a mensuales. Autores como Barcellan [2002] y Lupi y Parigi [2002] señalan que si se supone que los errores de alta frecuencia siguen una distribución normal, entonces se pueden estimar β , α y σ^2 utilizando el método de máxima verosimilitud. Estos autores mencionan un método alternativo para estimar α , el cual consiste en minimizar $\hat{\mu}'CVC'\hat{\mu} = (\mathbf{Y} - W\hat{\beta})'CVC'(\mathbf{Y} - W\hat{\beta})$, respecto a α . La estimación de α se hace a través de una búsqueda por rejilla (*grid search*) en el que se asignan valores de α entre -1 y 1 , y se escoge el valor que maximice la función de verosimilitud logarítmica.⁴

Aunque el método de Chow y Lin tiene un enfoque diferente al de Denton, las soluciones de ambos métodos tienen una estructura similar. Los dos métodos obtienen una solución que se puede expresar como una combinación lineal entre la discrepancia agregada y una *proxy* de la variable a estimar. Además, en ambos casos la solución depende de una función objetivo que se puede expresar como una forma cuadrática que depende de una matriz desconocida (A en el caso de Denton y V en el caso de Chow y Lin).

La principal desventaja del método de Chow y Lin, es que sólo suponen dos procesos muy simples para los errores de alta frecuencia, ruido blanco y AR(1).

3.3. Método de Fernández [1981]. Dada la relación que parece existir entre los métodos de Denton [1971] y Chow y Lin [1971], Fernández [1981] propone un método de desagregación que consiste en una reformulación de método de Denton en el contexto de estimación óptima tratado por Chow y Lin. En principio, Fernández generaliza el método de Denton, y encuentra los supuestos adicionales para que las soluciones de Denton y de Chow y Lin sean equivalentes. Lo anterior implica que la solución de Fernández se puede expresar como un estimador MELI de \mathbf{y} , y

⁴Véase Di-Fonzo [1987] para más detalles.

que además dicha estimación no daría lugar a una serie con discontinuidades artificiales.

Para generalizar el método de Denton, Fernández propone el siguiente problema de minimización:

$$\min_{\mathbf{y}} \rho(\mathbf{y}, w) = (\mathbf{y} - w\boldsymbol{\beta})'A(\mathbf{y} - w\boldsymbol{\beta}) \quad s.a. \quad C\mathbf{y} = \mathbf{Y} \quad (18)$$

Al igual que en los casos anteriores, la función de pérdida dada en (18) corresponde a una forma cuadrática. Como caso particular, cuando $\boldsymbol{\beta} = \mathbf{1}$, se obtiene la función objetivo de Denton. Fernández muestra que la solución a este problema de minimización está dada por:

$$\hat{\boldsymbol{\beta}} = [W'(CA^{-1}C')^{-1}W]^{-1}W'(CA^{-1}C')^{-1}\mathbf{Y} \quad (19)$$

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + A^{-1}C'(CA^{-1}C')^{-1}(\mathbf{Y} - W\hat{\boldsymbol{\beta}}) \quad (20)$$

Es importante notar que la expresión (20) es equivalente a (8) cuando $\boldsymbol{\beta} = \mathbf{1}$ y $w = \bar{y}$. Fernández también relaciona el enfoque de la función de pérdida con el marco teórico de la regresión lineal. Este autor señala que los dos enfoques generan el mismo resultado cuando A es la matriz identidad y los errores son homoscedásticos y no presentan correlación serial. Si $A = I_N$, la ecuación (18) corresponde a la suma de los errores al cuadrado, la cual equivale a la función objetivo del método de mínimos cuadrados usado en los modelos de regresión lineal. Si se mantienen los supuestos tradicionales sobre el término de error y $A = I_N$, la solución a (18) es un estimador MELI de \mathbf{y} .

No obstante, existe otra forma de relacionar los enfoques de minimización de función de pérdida cuadrática y MELI. Para mostrar este punto, Fernández supone que existe una relación lineal similar a (12), donde sus errores siguen una caminata aleatoria, es decir:

$$u_t = u_{t-1} + \varepsilon_t \quad (21)$$

Adicionalmente, se supone que las innovaciones de (21), ε_t , siguen un proceso ruido blanco con media cero y varianza σ^2 . También se asume que $u_0 = 0$, de tal forma que $Var(u_1) = \sigma^2$. Tomando primeras diferencias de (12) y utilizando la matriz D definida en (10), se obtiene el siguiente modelo:

$$D\mathbf{y} = Dw\boldsymbol{\beta} + D\mathbf{u} \quad (22)$$

donde $E(D\mathbf{u}\mathbf{u}'D') = \sigma^2 I_N$. Fernández anota que transformar el modelo (22) a baja frecuencia, es decir premultiplicándolo por C , no es útil porque la agregación de las primeras diferencias de las variables flujo no es una magnitud observable. Por esta razón, realiza una transformación adicional en la cual relaciona la primera diferencia en baja frecuencia con la primera diferencia en alta frecuencia, y muestra que el mejor estimador lineal de \mathbf{y} dados los supuestos anteriores es de la forma:

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + (D'D)^{-1}C'(C(D'D)^{-1}C')^{-1}(\mathbf{Y} - W\hat{\boldsymbol{\beta}}) \quad (23)$$

$$\hat{\boldsymbol{\beta}} = [W'(C(D'D)^{-1}C')^{-1}W]^{-1}W'(C(D'D)^{-1}C')^{-1}\mathbf{Y} \quad (24)$$

Las expresiones (20) y (23) son idénticas cuando $A = D'D$, lo que implica que en este caso, el enfoque de minimización de pérdida cuadrática tiene una solución que es equivalente al MELI de \mathbf{y} .

El método de Fernández proporciona un marco formal al enfoque de minimización de una función de pérdida planteado por Denton [1971], al incluir supuestos que lo relacionan con estimadores de un modelo de regresión lineal. Además, Fernández evita el problema de la estimación de parámetros adicionales (como α en el método de Chow y Lin), al suponer una caminata aleatoria en los errores.

Sin embargo, el supuesto de la caminata aleatoria sobre los errores de (12) implica que las series de alta frecuencia (\mathbf{y} y w) no pueden estar cointegradas, lo cual es una restricción importante en la práctica.

3.4. Método de Litterman [1983]. Litterman [1983] desarrolla un método que busca generalizar el propuesto por Fernández [1981]. Al igual que Fernández, Litterman supone la existencia de una relación lineal entre la serie a estimar y un conjunto de series indicadoras. Esta relación equivale a la ecuación (12) de la metodología de Chow y Lin.

Litterman propone que los errores de (12), siguen el proceso:

$$u_t = u_{t-1} + \varepsilon_t, \quad (25)$$

con

$$\varepsilon_t = \alpha \varepsilon_{t-1} + e_t \quad (26)$$

donde e_t es un proceso ruido blanco con varianza σ^2 . Es decir, Litterman conserva el supuesto de caminata aleatoria de Fernández, como se muestra en la ecuación (25). Sin embargo, no supone que el error de (12) es un proceso ruido blanco.

Las expresiones (25) y (26) implican:

$$e_t = u_t - (\alpha + 1)u_{t-1} + \alpha u_{t-2} \quad (27)$$

La expresión (27) se puede escribir en forma vectorial como:

$$\mathbf{e} = H\mathbf{D}\mathbf{u} \quad (28)$$

donde $\mathbf{e} = [e_1, \dots, e_N]$, D es la matriz definida en (10), y la matriz H tiene la siguiente forma:

$$H = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\alpha & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\alpha & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\alpha & 1 \end{bmatrix} \quad (29)$$

La ecuación (28) implica que el vector de errores de alta frecuencia u tiene la siguiente forma:

$$\mathbf{u} = D^{-1}H^{-1}\mathbf{e} \quad (30)$$

Por lo tanto, la varianza de \mathbf{u} se puede expresar como:

$$\text{Var}(\mathbf{u}) = V = (D'H'HD)^{-1}\sigma^2 \quad (31)$$

La solución de Litterman se obtiene al sustituir (31) en la solución de Chow y Lin dada en la expresión (14):

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + (D'H'HD)^{-1}C'(C(D'H'HD)^{-1}C')^{-1}(\mathbf{Y} - W\hat{\boldsymbol{\beta}}), \quad (32)$$

donde

$$\hat{\boldsymbol{\beta}} = [W'(C(D'H'HD)^{-1}C')^{-1}W]^{-1}W'(C(D'H'HD)C')^{-1}\mathbf{Y} \quad (33)$$

Las expresiones (32) y (33) dependen de la matriz H , que a su vez depende del parámetro α . Litterman propone un método para estimar α , pero sólo es aplicable a la desagregación de series trimestrales a mensuales.

Aunque Litterman generaliza el enfoque de Fernández [1981], éste no supera el problema principal del método Fernández ya que no admite la existencia de relaciones de cointegración entre y_t y w_t . Para probar lo anterior, observe que los supuestos sobre u_t implican que $u_t = y_t - w_t'\boldsymbol{\beta} \sim \text{ARIMA}(1,1,0)$ y, por lo tanto, $(1, -\boldsymbol{\beta}')$ no se puede interpretar como un vector de cointegración debido a que u_t es una serie integrada de orden uno. Por otro lado, el autor no desarrolla un método general⁵ para estimar el parámetro α .

3.5. Método de Guerrero [1990]. El método de Guerrero [1990] al igual que los métodos de Chow y Lin [1971], Fernández [1981] y Litterman [1983], parte de la existencia de una relación lineal de las variables de alta frecuencia. Además, su enfoque se centra en la estructura que sigue la variable y , y no en la de los errores de alta frecuencia de la ecuación (12). De la misma forma que el método de Denton [1971], el método de Guerrero [1990] requiere una estimación preliminar de la variable de interés con el fin de estimar la serie de alta frecuencia.

Para resolver el problema de desagregación, Guerrero supone que la serie de alta frecuencia no observada, y_t , sigue un proceso ARIMA de la forma:

$$\phi(B)(1-B)^d y_t = \tau(B)u_t \quad (34)$$

donde B es el operador de rezago, $\phi(B)$ y $\tau(B)$ son polinomios de rezagos de orden finito con coeficientes conocidos, d indica el grado de diferenciación de y_t , y u_t es un proceso ruido blanco con varianza σ_u^2 . Dada esta estructura de y_t , el estimador lineal de mínimo error cuadrático medio de y_t , $t > 0$, basado en información hasta el periodo 0, está dado por:

$$E(y_t|y_0, y_{-1}, \dots) = E_0(y_t) \quad (35)$$

El error de pronóstico en términos de la representación de media móvil pura es:

⁵Es decir, para cualquier tipo de frecuencia de los datos

$$y_t - E_0(y_t) = \sum_{i=0}^{t-1} \theta_i u_{t-i} \quad (36)$$

donde $\theta_0 \equiv 1$, y $\theta_1, \theta_2, \dots$ son los coeficientes del polinomio $\theta(B)$ tales que $\theta(B)\phi(B)(1-B)^d \tau^{-1}(B) = 1$. La representación matricial de la expresión (36) es:

$$\mathbf{y} - E_0(\mathbf{y}) = \boldsymbol{\theta} \mathbf{u} \quad (37)$$

donde $\boldsymbol{\theta}$ es una matriz triangular inferior formada por la secuencia $1, \theta_1, \dots, \theta_{N-1}$ en la primera columna, $0, 1, \theta_1, \dots, \theta_{N-2}$ en la segunda columna, y así sucesivamente. \mathbf{u} es un vector aleatorio con $E_0(\mathbf{u}) = \mathbf{0}$, y $E_0(\mathbf{u}\mathbf{u}') = \sigma_u^2 I_N$. Al incorporar el conjunto de restricciones de agregación, $\mathbf{Y} = C\mathbf{y}$, Guerrero muestra que la solución al problema de desagregación está dada por:

$$\hat{\mathbf{y}} = E_0(\mathbf{y}) + \boldsymbol{\theta}\boldsymbol{\theta}'C'(C\boldsymbol{\theta}\boldsymbol{\theta}'C')^{-1}(\mathbf{Y} - CE_0(\mathbf{y})) \quad (38)$$

En la práctica, se debe estimar $E_0(\mathbf{y})$. Para tal fin, Guerrero supone que existe una estimación preliminar de \mathbf{y} , $\tilde{\mathbf{y}}$, que debe satisfacer los siguientes supuestos:

1. \mathbf{y} y $\tilde{\mathbf{y}}$ tienen la misma estructura de autocorrelación, es decir, admiten el mismo modelo ARIMA,
2. $E(y_t|\tilde{\mathbf{y}}) = \tilde{y}_t$,
3. $E_0(\mathbf{y})$ es independiente de $\tilde{\mathbf{y}}$.

Dados estos supuestos, el valor esperado de la expresión (37), condicionado a $\tilde{\mathbf{y}}$, es:

$$\tilde{\mathbf{y}} - E_0(\mathbf{y}|\tilde{\mathbf{y}}) = E(\mathbf{y}|\tilde{\mathbf{y}}) - E[E_0(\mathbf{y})|\tilde{\mathbf{y}}] = \boldsymbol{\theta}E(\mathbf{u}|\tilde{\mathbf{y}}) \quad (39)$$

Substrayendo (39) de (37), se llega a la siguiente expresión:

$$\mathbf{y} - \tilde{\mathbf{y}} = \boldsymbol{\theta}\boldsymbol{\omega} \quad (40)$$

donde $\boldsymbol{\omega} = \mathbf{u} - E(\mathbf{u}|\tilde{\mathbf{y}})$ es un vector aleatorio con $E(\boldsymbol{\omega}|\tilde{\mathbf{y}}) = \mathbf{0}$ y $E(\boldsymbol{\omega}\boldsymbol{\omega}'|\tilde{\mathbf{y}}) = \sigma^2 P$, con P una matriz definida positiva. Como resultado del primer supuesto, y notando la similitud de las expresiones (37) y (40), un estimador MELI de \mathbf{y} está dado por:

$$\hat{\mathbf{y}} = \tilde{\mathbf{y}} + \boldsymbol{\theta}P\boldsymbol{\theta}'C'(C\boldsymbol{\theta}P\boldsymbol{\theta}'C')^{-1}(\mathbf{Y} - C\tilde{\mathbf{y}}) \quad (41)$$

con

$$Var[(\hat{\mathbf{y}} - \mathbf{y})|\tilde{\mathbf{y}}] = \sigma^2[I_N - \boldsymbol{\theta}P\boldsymbol{\theta}'C'(C\boldsymbol{\theta}P\boldsymbol{\theta}'C')^{-1}C]\boldsymbol{\theta}P\boldsymbol{\theta}' \quad (42)$$

Sin embargo, la estimación final de \mathbf{y} depende de σ^2 y P , que son desconocidos. Guerrero sugiere el siguiente procedimiento para estimar σ^2 y P :

1. Se supone $P = I_N$ y se calcula $\hat{\mathbf{y}}$ en la expresión (41). Posteriormente, se construye la serie $\hat{\boldsymbol{\omega}} = \boldsymbol{\theta}^{-1}(\hat{\mathbf{y}} - \tilde{\mathbf{y}})$. Si $\hat{\boldsymbol{\omega}}$ es ruido blanco, entonces el supuesto $P = I_N$ se acepta; por lo tanto, $\hat{\sigma}^2 = \hat{\boldsymbol{\omega}}\hat{\boldsymbol{\omega}}'/(N-r)$ es un estimador apropiado de σ^2 , donde r es el número de parámetros estimados en el modelo ARIMA de \tilde{y}_t . En el caso de que $\hat{\boldsymbol{\omega}}$ no sea ruido blanco, se debe construir un modelo ARIMA sobre $(\hat{y}_t - \tilde{y}_t)$. Con base en este resultado, se obtiene un

estimador de la matriz Ω que satisface la expresión $\mathbf{e} = Q\boldsymbol{\omega} = Q\boldsymbol{\theta}^{-1}(\mathbf{y} - \tilde{\mathbf{y}}) = \Omega^{-1}(\mathbf{y} - \tilde{\mathbf{y}})$, tal que $QPQ' = I_N$.

2. Se estima y a partir de (41) teniendo en cuenta que $\Omega\Omega' = \boldsymbol{\theta}(Q'Q)^{-1}\boldsymbol{\theta}' = \boldsymbol{\theta}P\boldsymbol{\theta}'$. Adicionalmente, se calcula la matriz de covarianzas dada en la expresión (42), donde el estimador de σ^2 está dado por $\hat{\sigma}^2 = \hat{\mathbf{e}}\hat{\mathbf{e}}'/(N-r) = (\hat{\mathbf{y}} - \tilde{\mathbf{y}})(\boldsymbol{\theta}P\boldsymbol{\theta}')^{-1}(\hat{\mathbf{y}} - \tilde{\mathbf{y}})/(N-r)$, con $N-r$ los grados de libertad del modelo ARIMA correspondiente.

En el método de Guerrero, al igual que en el de Denton, la estimación final de la serie desagregada depende de una estimación preliminar de dicha serie. Una ventaja de este método, a diferencia del de Denton, es que se proporciona una metodología para obtener $\tilde{\mathbf{y}}$. Guerrero propone una estimación preliminar a partir de un conjunto de series indicadoras, w , la cual tiene la siguiente forma:

$$\tilde{\mathbf{y}} = w\hat{\boldsymbol{\beta}} \quad (43)$$

Al sustituir (43) en la versión agregada de (40), se obtiene el siguiente resultado:

$$\mathbf{Y} = C\tilde{\mathbf{y}} + C\boldsymbol{\theta}\boldsymbol{\omega} = Cw\hat{\boldsymbol{\beta}} + C\boldsymbol{\theta}\boldsymbol{\omega} \quad (44)$$

La expresión anterior se puede notar de la siguiente forma:

$$\mathbf{Y} = W\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (45)$$

donde $\boldsymbol{\varepsilon} = C\boldsymbol{\theta}\boldsymbol{\omega}$. Por lo tanto, $\hat{\boldsymbol{\beta}}$ puede ser obtenido utilizando mínimos cuadrados ordinarios sobre (45).

Otra ventaja del método de Guerrero [1990], en comparación con los métodos de Chow y Lin [1971], Fernández [1981] y Litterman [1983], es que no requiere realizar supuestos sobre los errores de alta frecuencia.

Como lo señalan Lupi y Parigi [2002], una de las críticas al método de Guerrero es que éste no se puede aplicar cuando existe una relación de cointegración entre la serie preliminar y la serie no observada de alta frecuencia. Para ilustrar este punto, considere que $y_t \sim I(1)$ y $\tilde{y}_t = \mathbf{w}'_t\boldsymbol{\beta} \sim I(1)$. Dados los supuestos iniciales de Guerrero, se tiene que las dos series siguen la misma representación ARIMA, de modo que:

$$\phi(B)(1-B)y_t = \theta(B)\varepsilon_{1t} \quad (46)$$

$$\phi(B)(1-B)\tilde{y}_t = \theta(B)\varepsilon_{2t} \quad (47)$$

donde $\phi(B)$ y $\theta(B)$ son dos polinomios de rezagos de orden finito, cuyos coeficientes son conocidos, y ε_{1t} y ε_{2t} son dos procesos ruido blanco ortogonales con varianzas σ_1^2 y σ_2^2 , respectivamente. Al realizar una combinación lineal entre (46) de (47), se tiene la siguiente expresión:

$$(1-B)\phi(B)(y_t + \gamma\tilde{y}_t) = \theta(B)\eta_t \quad (48)$$

donde $\eta_t = \varepsilon_{1t} + \gamma\varepsilon_{2t}$. De (48) se puede concluir que la combinación lineal entre las series desagregadas no es estacionaria y, por lo tanto, las dos series no pueden estar cointegradas.

Por otro lado, si se supone que y_t y $\tilde{y}_t = \mathbf{w}'_t\boldsymbol{\beta}$ son generados por el siguiente proceso bivariado:

$$\begin{bmatrix} 1 & -\delta \\ 0 & (1-B) \end{bmatrix} \begin{bmatrix} y_t \\ \tilde{y}_t \end{bmatrix} = \begin{bmatrix} \theta_{11}(B) & 0 \\ 0 & \theta_{22}(B) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (49)$$

donde $\theta_{11}(B)$ y $\theta_{22}(B)$ son dos polinomios de rezagos con coeficientes conocidos. La primera ecuación del sistema (49) implica que las dos series están cointegradas, donde el vector de cointegración es $[1, -\delta]'$, mientras que la segunda ecuación implica que \tilde{y}_t es I(1).

Se puede mostrar que si $\theta_{11}(B) \equiv 1$ y $\delta \equiv 1$, entonces $\Delta y_t = \Delta\varepsilon_{1t} + \theta_{22}(B)\varepsilon_{2t}$, lo que implica que y_t y \tilde{y}_t , no pueden ser generados por el mismo proceso ARIMA. En resumen, se puede mostrar que si la serie preliminar y la variable de interés siguen el mismo proceso ARIMA no pueden estar cointegradas, y que si se parte de que están cointegradas no necesariamente siguen el mismo proceso ARIMA.

3.6. Método de Santos-Silva y Cardoso [2001]. Al igual que Guerrero [1990], Santos-Silva y Cardoso [2001] proponen un método que supone una estructura sobre las variables de alta frecuencia, y no sobre los errores asociados a la relación lineal entre dichas variables. A diferencia de los métodos de Chow y Lin [1971], Fernández [1981] y Litterman [1983] que proponen un modelo lineal estático para las series de alta frecuencia, ecuación (12), Santos-Silva y Cardoso [2001] proponen un modelo lineal dinámico de la siguiente forma:

$$y_t = ky_{t-1} + w_t\gamma + v_t \quad (50)$$

donde $|k| \leq 1$ y v_t es un proceso ruido blanco. Como lo señalan Santos Silva y Cardoso, este método puede ser utilizado incluyendo rezagos adicionales de y_t y rezagos de las variables indicadoras en el vector \mathbf{w}_t . Asumiendo $|k| < 1$, y realizando sustituciones recursivas, el modelo (50) puede escribirse como:

$$y_t = \left(\sum_{i=0}^{\infty} k^i w_{t-i} \right) \gamma + \left(\sum_{i=0}^{\infty} k^i v_{t-i} \right) \quad (51)$$

La ecuación (51) también puede expresarse como:

$$y_t = w(k)_t \gamma + k^t \eta + u_t \quad (52)$$

donde $w(k)_t = \sum_{i=0}^{t-1} k^i w_{t-i}$, $\eta = \sum_{i=0}^{\infty} k^i w_{-i} \gamma$ y $u_t = \sum_{i=0}^{\infty} k^i v_{t-i} = kv_{t-1} + v_t$ es un error estacionario que sigue un proceso AR(1). Realizando las agrupaciones apropiadas, la ecuación (52) se puede expresar en forma vectorial como:

$$y = \mathbf{x}(k)\boldsymbol{\beta} + u, \quad (53)$$

con $\mathbf{x} = [\mathbf{w}(k) \quad K]$, $\mathbf{w}(k) = [w(k)_0, \dots, w(k)_N]'$, $K = [1, k, k^2, \dots, k^N]'$ y $\beta = [\gamma, \eta]'$. Premultiplicando (53) por la matriz de agregación \mathbf{C} , se obtiene:

$$Y = \mathbf{C}\mathbf{x}(k)\beta + \mathbf{C}u \quad (54)$$

Nótese que y depende de los parámetros k , η y γ . Como se puede apreciar en la ecuación (52), esta dependencia es lineal en η y γ , pero no en k . Por consiguiente, si se parte de un estimador de k , \hat{k} , los estimadores de η y γ pueden ser obtenidos mediante el método de mínimos cuadrados generalizados sobre la ecuación (55). Adicionalmente, Santos Silva y Cardoso muestran que un estimador de la serie desagregada es el siguiente:

$$\hat{y} = \mathbf{x}(\hat{k})\hat{\beta}(\hat{k}) + \Omega(\hat{k})\mathbf{C}'(\mathbf{C}\Omega(\hat{k})\mathbf{C}')^{-1}(Y - \mathbf{C}\mathbf{x}(\hat{k})\hat{\beta}(\hat{k})) \quad (55)$$

donde $\hat{\beta}(\hat{k}) = [\mathbf{x}(\hat{k})'\mathbf{C}'(\mathbf{C}\Omega(\hat{k})\mathbf{C}')^{-1}\mathbf{C}\mathbf{x}(\hat{k})]^{-1}\mathbf{x}(\hat{k})'\mathbf{C}'(\mathbf{C}\Omega(\hat{k})\mathbf{C}')^{-1}Y$, y $\Omega(\hat{k})$ es la matriz de covarianzas de u condicionada a \hat{k} , con $u_0 = 0$. Los autores sugieren estimar k mediante el método de máxima verosimilitud.

Una de las ventajas del método de Santos Silva y Cardoso, es que supone una estructura sobre las series desagregadas y no sobre los errores asociados a los modelos de alta frecuencia. Adicionalmente, en contraste al método de Guerrero, este método es válido cuando existen relaciones de cointegración entre las variables de alta frecuencia. Los autores afirman que el modelo dinámico que ellos proponen se puede expresar como un modelo de corrección de errores.

3.7. Desagregación temporal en caso de transformaciones logarítmicas. En algunas ocasiones, se requiere transformar las series utilizando funciones de potencia. Una transformación convencional en este contexto es la función logaritmo. Esta transformación puede generar dificultades cuando se aplica a ciertos problemas de desagregación, en particular cuando la variable a desagregar es un flujo debido a que la función logaritmo no es aditiva. Es decir, si $Y_s = \sum_{m=1}^M y_{m,s}$ entonces $\ln Y_s \neq \sum_{m=1}^M \ln y_{m,s}$.

Con el fin de superar el problema descrito anteriormente, Aadland [2000] propuso un método para linealizar la restricción de agregación. Este método supone que el modelo lineal que relaciona la serie desagregada con las variables indicadoras tiene la siguiente forma:

$$\ln y_{m,s} = \mathbf{w}'_{m,s}\beta + u_{m,s} \quad (56)$$

donde $u_{m,s}$ es un proceso ruido blanco con $E(u_{\tilde{m},\tilde{s}}\mathbf{w}_{m,s}) = \mathbf{0}$ para todo \tilde{m} , \tilde{s} , m y s .

Para linealizar $\ln y_{m,s}$, este autor utiliza la aproximación de Taylor de primer orden alrededor de \bar{Y}_s :

$$\ln y_{m,s} \cong \ln \bar{Y}_s + \frac{1}{\bar{Y}_s} (y_{m,s} - \bar{Y}_s) = \ln Y_s - \ln M + \frac{M y_{m,s}}{Y_s} - 1 \quad (57)$$

donde $\bar{Y}_s = \frac{1}{M} \sum_{m=1}^M y_{m,s} = \frac{Y_s}{M}$ y $Y_s = \sum_{m=1}^M y_{m,s}$.

La suma de (57) sobre los periodos de alta frecuencia está dada por:

$$Z_s = \sum_{m=1}^M z_{m,s} \cong M \ln Y_s - M \ln M + \frac{M \sum_{m=1}^M y_{m,s}}{Y_s} - M = M \ln Y_s - M \ln M \quad (58)$$

donde $z_{m,s} \equiv \ln y_{m,s}$.

A partir de la expresión anterior, se puede definir el siguiente modelo agregado en términos de variables observables:

$$Z_s = \mathbf{W}'_s \boldsymbol{\beta} + u_s \quad (59)$$

donde $Z_s = M \ln Y_s - M \ln M$, $\mathbf{W}_s = \sum_{m=1}^M w_{m,s}$ y $u_s = \sum_{m=1}^M u_{m,s}$. Un estimador de $z_{m,s}$, $\hat{z}_{m,s}$, puede ser obtenido aplicando cualquier método de desagregación con base en la ecuación (59). Finalmente, una estimación de la serie desagregada en niveles puede ser obtenida como $\hat{y}_{m,s} = \exp(\hat{z}_{m,s})$.

Debido a la aproximación de Taylor dada en (57), $\hat{y}_{m,s}$ no cumple exactamente las restricciones de agregación. Para solucionar este problema, Proietti [1999] sugiere tomar la estimación resultante, $\hat{y}_{m,s}$ como la estimación preliminar del método de Denton [1971], y obtener una nueva estimación de $y_{m,s}$ a partir de dicho método.

4. MÉTODOS MULTIVARIADOS

En esta sección se hace una reseña de algunos de los métodos multivariados de desagregación temporal y se propone una extensión del método de Di-Fonzo [1990]. En la Tabla 2 se resumen las características principales de estos métodos.

4.1. Método de Rossi [1982]. Rossi [1982] generaliza el método de Chow y Lin [1971] al caso multivariado. Este método parte de una matriz de dimensión $S \times H$ de series agregadas $Y = [\mathbf{Y}_1, \dots, \mathbf{Y}_H]$, donde \mathbf{Y}_h es un vector de dimensión $S \times 1$ que contiene los valores de baja frecuencia de la serie h , con $h = 1, \dots, H$. El problema consiste en estimar la matriz de dimensión $N \times H$, $y = [y_1, \dots, y_H]$, que contiene las series de alta frecuencia correspondientes. Al igual que las metodologías descritas en las secciones anteriores, Rossi supone que existe la siguiente relación:

$$\mathbf{y} = w\boldsymbol{\beta} + \mathbf{u} \quad (60)$$

donde $\mathbf{y} = \text{vec}(y)$ es un vector de dimensión $NH \times 1$ obtenido al aplicar el operador $\text{vec}(\cdot)$ a la matriz y .⁶ La matriz w es una matriz diagonal en bloques de dimensión $NH \times p$ definida de la siguiente forma:

$$w = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_H \end{bmatrix} \quad (61)$$

⁶El operador $\text{vec}(\cdot)$ convierte una matriz en un vector columna al tomar las columnas de la matriz y ponerlas una debajo de la otra.

Método	Relación de alta frecuencia	Supuestos sobre el error	Desventajas
Rossi [1982]	<ul style="list-style-type: none"> • $\mathbf{y} = \mathbf{w}\boldsymbol{\beta} + \mathbf{u}$ 	$\mathbf{u} \sim$ ruido blanco multivariado	<ul style="list-style-type: none"> • El supuesto sobre el error es demasiado simplista • No siempre se satisfacen las restricciones de agregación temporales
Di-Fonzo [1990]	<ul style="list-style-type: none"> • $\mathbf{y} = \mathbf{w}\boldsymbol{\beta} + \mathbf{u}$ • La ecuación agregada incluye restricciones contemporáneas y temporales 	$\mathbf{u} \sim$ ruido blanco multivariado	<ul style="list-style-type: none"> • El supuesto sobre el error es demasiado simplista
Propuesto	<ul style="list-style-type: none"> • $\mathbf{y} = \mathbf{w}\boldsymbol{\beta} + \mathbf{u}$ • La ecuación agregada incluye restricciones contemporáneas y temporales 	$\mathbf{u} \sim VAR(1)$	<ul style="list-style-type: none"> • El supuesto sobre el error es simplista
Moauero y Savio [2005]	<ul style="list-style-type: none"> • Modelo tipo SUTSE 	—	<ul style="list-style-type: none"> • Se requieren los valores iniciales del vector de estado y de la matriz de error cuadrático medio

En estos métodos $\mathbf{y} = \text{vec}(y)$ es un vector de dimensión $NH \times 1$, donde N es el número de observaciones en el tiempo y H el número de series a desagregar, $y = [y_1, \dots, y_H]$, $\mathbf{y}_h = [y_{h1}, \dots, y_{hN}]'$ corresponde a los valores de la h -ésima serie de alta frecuencia que se desea estimar, $h = 1, \dots, H$; Y_s son los valores observados de baja frecuencia, con $s = 1, \dots, S$; \mathbf{w} corresponde a los valores de las series indicadoras en alta frecuencia asociadas a \mathbf{y} . Finalmente, \mathbf{u} son los errores de alta frecuencia

TABLA 2. Principales características de algunos métodos multivariados de desagregación temporal.

donde w_h es una matriz $N \times p_h$ de variables indicadoras asociadas a la variable h con $h = 1, \dots, H$. Por otro lado, $\boldsymbol{\beta}$ es un vector de parámetros de dimensión $p \times 1$ con $p = \sum_{h=1}^H p_h$ y $\mathbf{u} = \text{vec}(U) = \text{vec}[\mathbf{u}_1, \dots, \mathbf{u}_H]$ es un proceso ruido blanco multivariado con $E(\mathbf{u}\mathbf{u}') = \Sigma_U \otimes I_N$, donde Σ_U es una matriz definida positiva de dimensión $H \times H$.

Debido a su carácter multivariado, este método tiene la ventaja de considerar dos tipos de restricciones de agregación. La primera corresponde a la restricción de agregación temporal, similar a la expuesta en los métodos univariados. La segunda se relaciona con una restricción de agregación contemporánea.

Rossi define la agregación contemporánea de \mathbf{y} como el vector $\mathbf{z} = \mathbf{y}\mathbf{j}_H$, donde \mathbf{j}_H es un vector de unos de dimensión $H \times 1$. El método supone que \mathbf{z} es conocido. Si se define $\mathcal{B} = \mathbf{j}_H' \otimes I_N$ como la matriz de agregación contemporánea, entonces la ecuación que resulta de premultiplicar (60) por \mathcal{B} es:

$$\mathbf{z} = \mathcal{B}\mathbf{w}\boldsymbol{\beta} + \mathcal{B}\mathbf{u} \quad (62)$$

Rossi muestra que el mejor estimador lineal insesgado de \mathbf{y} está dado por:

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + (\Sigma_U \otimes I_N)\mathcal{B}'(\mathcal{B}(\Sigma_U \otimes I_N)\mathcal{B}')^{-1}(\mathbf{z} - \mathcal{B}w\boldsymbol{\beta}) \quad (63)$$

$$\hat{\boldsymbol{\beta}} = (w'\mathcal{B}'(\mathcal{B}(\Sigma_U \otimes I_N)\mathcal{B}')^{-1}\mathcal{B}w)^{-1}w'\mathcal{B}'(\mathcal{B}(\Sigma_U \otimes I_N)\mathcal{B}')^{-1}\mathbf{z} \quad (64)$$

Es importante anotar que $\hat{\mathbf{y}}$ satisface la restricción de agregación contemporánea, es decir, $\mathbf{z} = \hat{\mathbf{y}}\mathbf{j}_H$. Sin embargo, no se garantiza que las restricciones de agregación temporal se satisfagan. Para imponer estas restricciones, se define la matriz de agregación temporal $\mathcal{C} = (I_H \otimes C)$, donde C es la matriz de agregación temporal definida en (2). De esta forma, el conjunto de restricciones de agregación temporales del caso multivariado está dado por $\mathcal{C}\mathbf{y} = \mathbf{Y}$, donde $\mathbf{Y} = \text{vec}(Y)$. En este caso, se puede probar que no se satisface la igualdad $\mathcal{C}\hat{\mathbf{y}} = \mathbf{Y}$. Por esta razón, Rossi redefine la matriz w tal que:

$$w = \begin{bmatrix} \tilde{\mathbf{y}}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{y}}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathbf{y}}_H \end{bmatrix} \quad (65)$$

donde $\tilde{\mathbf{y}}_h$ es el mejor estimador lineal insesgado de \mathbf{y}_h , dado en (14), aplicando el método univariado de Chow y Lin [1971]⁷ para la variable h y sus variables indicadoras asociadas w_h , con $h = 1, \dots, H$. Si se reemplaza la matriz w definida en (65) en (63) y (64), se puede probar que $\mathcal{C}\hat{\mathbf{y}} = \mathbf{Y}$, es decir, que se garantiza que el estimador satisface tanto la restricción de agregación contemporánea como las restricciones temporales.

Para encontrar la matriz Σ_U , se debe suponer alguna estructura en los errores de (60). El caso más sencillo es $\Sigma_U = \sigma_U^2 I_H$, lo que implica que $(\Sigma_U \otimes I_N)\mathcal{B}'(\mathcal{B}(\Sigma_U \otimes I_N)\mathcal{B}')^{-1} = (1/H)\mathcal{B}'$. Por lo tanto, $\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + (1/H)\mathcal{B}'(\mathbf{z} - \mathcal{B}w\hat{\boldsymbol{\beta}})$, es decir, la discrepancia $\mathbf{z} - \mathcal{B}'w\hat{\boldsymbol{\beta}}$ se reparte en partes iguales entre las H variables.

Rossi describe un caso alternativo en el cual la matriz Σ_U es diagonal, donde el elemento hh de esta matriz, $\sigma_{U, hh}$, se supone proporcional a la media del indicador h , \bar{y}_h . En este caso, $(\Sigma_U \otimes I_N)\mathcal{B}'(\mathcal{B}(\Sigma_U \otimes I_N)\mathcal{B}')^{-1} = (\bar{\mathbf{w}} \otimes I_N)$, donde $\bar{\mathbf{w}} = [\bar{y}_1, \dots, \bar{y}_H]'$, lo que significa que la discrepancia se distribuye de forma proporcional o *pro rata*.

Una de las desventajas del método de Rossi es que no considera el caso de extrapolación, limitando su uso en la práctica. Otra desventaja de este método es que al suponer que $E(\mathbf{u}\mathbf{u}') = \Sigma_U \otimes I_N$ se tiene que $E(\mathbf{u}_h\mathbf{u}_k') = \sigma_{hk}I_N$ con $h, k = 1, \dots, H$. Esta restricción constituye un supuesto fuerte ya que implica que no pueden existir estructuras de heteroscedasticidad ni autocorrelación en el tiempo.

Di-Fonzo [1990] muestra que $E(\hat{\boldsymbol{\beta}}) \neq \mathbf{j}_H$, dado que los regresores del modelo (62) son estocásticos.⁸ Lo anterior implica que $\hat{\mathbf{y}}$ será consistente con \mathbf{z} pero no con \mathbf{Y} , es decir, no se cumplen las

⁷De forma similar, $\tilde{\mathbf{y}}_h$ puede ser obtenido a través de cualquier método univariado expuesto en la Sección 3.

⁸Estos regresores son estocásticos debido a la definición de w en (65).

restricciones de agregación temporales ya que se tiene $\mathcal{C}\hat{\mathbf{y}} \neq \mathbf{Y}$. Bajo el esquema propuesto por Rossi, la única forma posible para que se cumplan las restricciones contemporáneas y temporales simultáneamente, es reemplazando $\boldsymbol{\beta} = \mathbf{j}_H$ en (64). Este resultado se puede interpretar como una generalización del método univariado de Denton [1971] al caso multivariado.

4.2. Método de Di-Fonzo [1990]. Di-Fonzo [1990] desarrolló un método que permite desagregar más de una serie simultáneamente. Esta metodología supone que se cuenta con H series agregadas \mathbf{Y}_h , $h = 1, \dots, H$, cada una de dimensión $S \times 1$. El problema consiste en estimar los vectores correspondientes de alta frecuencia \mathbf{y}_h , $h = 1, \dots, H$, de dimensión $N \times 1$, con $N = M S$, donde M son los subperiodos de cada periodo. Al igual que algunos métodos univariados, se supone que existe una relación lineal entre las variables desagregadas y un conjunto de variables indicadoras:

$$\mathbf{y}_h = w_h \boldsymbol{\beta}_h + \mathbf{u}_h \quad h = 1, \dots, H \quad (66)$$

donde w_h es una matriz de dimensión $N \times p_h$ que contiene las variables indicadoras asociadas a la serie h , $\boldsymbol{\beta}_h$ es un vector de dimensión $p_h \times 1$ de coeficientes desconocidos y \mathbf{u}_h es un vector aleatorio con media cero y $E(u_k u_h') = V_{kh}$, con $k, h = 1, \dots, H$. Como caso particular, cuando $H = 1$ la solución al problema de desagregación se obtiene a partir de las metodologías univariadas descritas en la sección anterior.

El sistema de ecuaciones descrito en (66) puede escribirse en forma compacta como:

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_H \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_H \end{bmatrix}}_w \underbrace{\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_H \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_H \end{bmatrix}}_{\mathbf{u}} \quad (67)$$

Es decir:

$$\mathbf{y} = w\boldsymbol{\beta} + \mathbf{u} \quad (68)$$

donde \mathbf{y} es el vector de valores de alta frecuencia no observados de dimensión $NH \times 1$, w es una matriz diagonal en bloques de variables indicadoras observadas en alta frecuencia de dimensión $NH \times p$, con $p = \sum_{h=1}^H p_h$, $\boldsymbol{\beta}$ es un vector de parámetros de dimensión $p \times 1$, y \mathbf{u} es un vector aleatorio de dimensión $NH \times 1$ con media cero y matriz de covarianzas V , donde el bloque kh de V está dado por V_{kh} , $k, h = 1, \dots, H$.

Di Fonzo define la restricción de agregación contemporánea como:

$$\sum_{h=1}^H \mathbf{y}_h = \mathbf{z} \quad (69)$$

Un ejemplo de este tipo de restricciones es discutido en la sección 5. De forma similar, esta restricción se puede escribir como:

$$(\mathbf{j}'_H \otimes I_N) \mathbf{y} = \mathbf{z} \quad (70)$$

con \mathbf{j}_H un vector de unos de dimensión $H \times 1$.

Por otro lado, el conjunto de restricciones de agregación temporal está dado por $\mathcal{C}\mathbf{y} = \mathbf{Y}$, donde $\mathbf{Y} = [\mathbf{Y}_1, \dots, \mathbf{Y}_H]'$ y $\mathcal{C} = I_H \otimes C$, donde C es la matriz de agregación temporal dada en (4). De esta forma, el conjunto completo de restricciones de agregación, temporal y contemporánea, está dado por:

$$B\mathbf{y} = \mathbf{Y}_a \quad (71)$$

En la expresión anterior $B = [\mathbf{j}_H \otimes I_N | \mathcal{C}]'$ y $\mathbf{Y}_a = [\mathbf{z}' | \mathbf{Y}']'$. Premultiplicando (67) por la matriz B , se tiene:

$$\underbrace{\begin{bmatrix} \mathbf{z} \\ \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_H \end{bmatrix}}_{\mathbf{Y}_a} = \underbrace{\begin{bmatrix} w_1 & w_2 & \cdots & w_H \\ W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_H \end{bmatrix}}_{W_a} \underbrace{\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \\ \vdots \\ \boldsymbol{\beta}_H \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \boldsymbol{\mu}_z \\ \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_H \end{bmatrix}}_{\boldsymbol{\mu}_a} \quad (72)$$

donde $W_h = Cw_h$ son matrices de variables indicadoras agregadas temporalmente de dimensión $S \times p_h$ y $\boldsymbol{\mu}_h = C\mathbf{u}_h$ son vectores de innovaciones de dimensión $S \times 1$, con $h = 1, \dots, H$. Adicionalmente, $\boldsymbol{\mu}_z = \sum_{h=1}^H \mathbf{u}_h$ es el vector de las innovaciones agregadas contemporáneamente. El modelo (72) puede reescribirse como $\mathbf{Y}_a = W_a\boldsymbol{\beta} + \boldsymbol{\mu}_a$, donde $W_a = Bw$ y $\boldsymbol{\mu}_a = B\mathbf{u}$ es un vector aleatorio con matriz de covarianzas de la forma $V_a = BV B'$, la cual es singular debido a que existe colinealidad entre la restricción contemporánea y las temporales.

Di Fonzo muestra que para los problemas de distribución e interpolación, el mejor estimador lineal insesgado de \mathbf{y} está dado por:

$$\hat{\mathbf{y}} = w\hat{\boldsymbol{\beta}} + VB'V_a^+(\mathbf{Y}_a - W_a\hat{\boldsymbol{\beta}}) \quad (73)$$

$$\hat{\boldsymbol{\beta}} = (W_a'V_a^+W_a)^{-1}W_a'V_a^+\mathbf{Y}_a \quad (74)$$

donde V_a^+ es la inversa generalizada de Moore-Penrose de V_a . Adicionalmente, la matriz de covarianzas del error es:

$$E[(\hat{\mathbf{y}} - \mathbf{y})(\hat{\mathbf{y}} - \mathbf{y})'] = (I_{NH} - VB'V_a^+B)V + (w - VB'V_a^+W_a)(W_a'V_a^+W_a)^{-1}(w - VB'V_a^+W_a)' \quad (75)$$

Al igual que en el caso univariado, la estimación de las series desagregadas depende de la forma de la matriz de varianzas y covarianzas de los errores de alta frecuencia, V . En este contexto, Di Fonzo supone que los errores de alta frecuencia son ruido blanco. Para simplificar el proceso de estimación, se redefine la matriz B de la siguiente forma:

$$B = \begin{bmatrix} \mathbf{j}'_H \otimes I_N \\ I_{H-1} \otimes C \quad \vdots \quad 0 \\ \hline 0 \quad \vdots \quad C \end{bmatrix} = \begin{bmatrix} B_r \\ B_S \end{bmatrix} \quad (76)$$

Di Fonzo muestra que la matriz B tiene rango $r = N + S(H - 1)$, y define las matrices P y R de dimensión $S \times r$ y $(r + S) \times r$, respectivamente, como: $P = [C \mid -(\mathbf{j}'_{H-1} \otimes I_S)]$ y $R = [I_r \mid P']'$. De esta forma, se tiene que $B_S = PB_r$ y $B = RB_r$. Adicionalmente, se puede mostrar que:

$$V_a = \begin{bmatrix} V_r & \vdots & V_r P' \\ P V_r & \vdots & P V_r P' \end{bmatrix} = [R V_r \mid R V_r P'] = R V_r R' \quad (77)$$

donde $V_r = E(\boldsymbol{\mu}_r \boldsymbol{\mu}'_r) = B_r V B'_r$ es una matriz de covarianzas $r \times r$ de rango completo con $\boldsymbol{\mu}_r = B_r \mathbf{u}$.

Incorporando las definiciones anteriores, los estimadores $\hat{\mathbf{y}}$ y $\hat{\boldsymbol{\beta}}$ descritos en (73) y (74) se pueden escribir como:

$$\hat{\mathbf{y}} = w \hat{\boldsymbol{\beta}} + V B'_r V_r^{-1} (\mathbf{Y}_r - W_r \hat{\boldsymbol{\beta}}) \quad (78)$$

$$\hat{\boldsymbol{\beta}} = (W'_r V_r^{-1} W_r)^{-1} W'_r V_r^{-1} \mathbf{Y}_r \quad (79)$$

Con $\mathbf{Y}_r = B_r \mathbf{Y}$ y $W_r = B_r w_r$. Cuando los errores \mathbf{u}_t son ruido blanco se tiene que el bloque kh de V , V_{kh} , está dado por $\sigma_{kh} I_N$ para $h, k = 1, \dots, H$. Di Fonzo muestra que:

$$V_r = \begin{bmatrix} \sigma I_N & \vdots & \boldsymbol{\sigma}'_r \otimes C' \\ \boldsymbol{\sigma}_r \otimes C & \vdots & (\mathbf{c}' \mathbf{c} \Sigma_r \otimes I_S) \end{bmatrix} \quad (80)$$

donde \mathbf{c} está definido como en (1), $\boldsymbol{\sigma} = \mathbf{j}'_H \Sigma \mathbf{j}_H$, $\boldsymbol{\sigma}_r$ es el vector $(H - 1) \times 1$ obtenido al eliminar el último elemento de $\Sigma \mathbf{j}_H$ y Σ_r es la matriz $(H - 1) \times (H - 1)$ obtenida al eliminar la última fila y la última columna de Σ . La matriz Σ , de dimensión $H \times H$ está definida de tal forma que su elemento kh es σ_{kh} .

Por último, si Σ es desconocida, ésta se estima como la matriz de varianzas y covarianzas de $\hat{\mathbf{U}}_1, \dots, \hat{\mathbf{U}}_H$, donde $\hat{\mathbf{U}}_h$ son los residuos de MCO de la regresión del modelo en baja frecuencia $\mathbf{Y}_h = W_h \boldsymbol{\beta}_h + \mathbf{U}_h$, para $h = 1, \dots, H$.

4.3. Método propuesto. En esta sección se propone una extensión del método de Di-Fonzo [1990]. Como se comentó anteriormente, el método de Di-Fonzo asume que los errores de alta frecuencia de la ecuación (68) son ruido blanco, el cual es un supuesto fuerte. En analogía a las metodologías de desagregación univariadas, un supuesto menos exigente corresponde a un modelo autorregresivo de orden uno sobre estos errores. En este contexto se puede asumir que estos errores siguen un modelo $VAR(1)$ definido de la siguiente forma:

$$\mathbf{u}_t^\bullet = A_1 \mathbf{u}_{t-1}^\bullet + \boldsymbol{\varepsilon}_t^\bullet, \quad t = 1, \dots, N \quad (81)$$

donde el superíndice “•” indica que los componentes de los vectores corresponden a diferentes variables en el mismo periodo de tiempo.⁹ En el caso de un VAR diagonal, la expresión (81) se puede reescribir de la siguiente forma:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \\ \vdots \\ u_{H,t} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{HH} \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \\ \vdots \\ u_{H,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{H,t} \end{bmatrix} \quad (82)$$

Los errores de la ecuación (81), ε_t^\bullet , siguen un proceso ruido blanco multivariado con matriz de varianzas y covarianzas, Γ_0 , de dimensión $H \times H$. Por otro lado, $-1 < a_{hh} < 1$ para $h = 1, \dots, H$, lo que implica un proceso estacionario.

En el Apéndice A se muestra que cuando los errores de la ecuación (68) siguen un VAR(1) diagonal, su matriz de covarianzas V tiene la siguiente forma:

$$V = \begin{bmatrix} \gamma_{11}^{(0)} A_1^{(11)} & \gamma_{12}^{(0)} A_1^{(01)} + \gamma_{21}^{(0)} A_1^{(20)} & \cdots & \gamma_{1H}^{(0)} A_1^{(01)} + \gamma_{H1}^{(0)} A_1^{(H0)} \\ \gamma_{21}^{(0)} A_1^{(02)} + \gamma_{12}^{(0)} A_1^{(10)} & \gamma_{22}^{(0)} A_1^{(22)} & \cdots & \gamma_{2H}^{(0)} A_1^{(02)} + \gamma_{H2}^{(0)} A_1^{(H0)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{H1}^{(0)} A_1^{(0H)} + \gamma_{1H}^{(0)} A_1^{(10)} & \gamma_{H2}^{(0)} A_1^{(0H)} + \gamma_{2H}^{(0)} A_1^{(20)} & \cdots & \gamma_{HH}^{(0)} A_1^{(HH)} \end{bmatrix} \quad (83)$$

donde:

$$A_1^{(hh)} = \begin{bmatrix} 1 & a_{hh} & a_{hh}^2 & \cdots & a_{hh}^{N-1} \\ a_{hh} & 1 & a_{hh} & \cdots & a_{hh}^{N-2} \\ a_{hh}^2 & a_{hh} & 1 & \cdots & a_{hh}^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{hh}^{N-1} & a_{hh}^{N-2} & a_{hh}^{N-3} & \cdots & 1 \end{bmatrix} \quad (84)$$

⁹En las ecuaciones anteriores, por ejemplo en (66), los componentes de los vectores corresponden a diferentes periodos de tiempo para la misma variable.

$$A_1^{(0h)} = \begin{bmatrix} 1 & a_{hh} & a_{hh}^2 & \cdots & a_{hh}^{N-1} \\ 0 & 1 & a_{hh} & \cdots & a_{hh}^{N-2} \\ 0 & 0 & 1 & \cdots & a_{hh}^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad (85)$$

$$A_1^{(h0)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{hh} & 0 & 0 & \cdots & 0 \\ a_{hh}^2 & a_{hh} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{hh}^{N-1} & a_{hh}^{N-2} & a_{hh}^{N-3} & \cdots & 0 \end{bmatrix} \quad (86)$$

para $h = 1, \dots, H$. Por su parte, para un modelo VAR(1) diagonal la matriz Γ_0 tiene la forma:

$$\Gamma_0 = \begin{pmatrix} \frac{\sigma_{11}^\varepsilon}{1-a_{11}^2} & \frac{\sigma_{12}^\varepsilon}{1-a_{11}a_{22}} & \cdots & \frac{\sigma_{1H}^\varepsilon}{1-a_{11}a_{HH}} \\ \frac{\sigma_{21}^\varepsilon}{1-a_{11}a_{22}} & \frac{\sigma_{22}^\varepsilon}{1-a_{22}^2} & \cdots & \frac{\sigma_{2H}^\varepsilon}{1-a_{22}a_{HH}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{H1}^\varepsilon}{1-a_{11}a_{HH}} & \frac{\sigma_{H2}^\varepsilon}{1-a_{22}a_{HH}} & \cdots & \frac{\sigma_{HH}^\varepsilon}{1-a_{HH}^2} \end{pmatrix} \quad (87)$$

donde σ_{kh}^ε es el elemento kh de la matriz de varianzas y covarianzas contemporáneas de $\boldsymbol{\varepsilon}_t, \Sigma^\varepsilon$. De esta forma, la matriz V únicamente depende de los parámetros a_{hh} y σ_{kh}^ε para $k, h = 1, \dots, H$.

En resumen, el objetivo de la desagregación es obtener $\hat{\mathbf{y}}$ utilizando las ecuaciones (73) y (74). Estas dependen de los datos observados y de la matriz V . Si se supone que los errores de la ecuación (68) siguen un modelo VAR(1) diagonal, la matriz V dependerá a su vez de las matrices A_1 y Σ^ε , y por consiguiente de los parámetros a_{hh} y σ_{kh}^ε para $k, h = 1, \dots, H$. Esto implica que para obtener una estimación de las series desagregadas, bajo un VAR(1), se debe contar con una estimación de los parámetros de las matrices A_1 y Σ^ε .

En el caso univariado Chow y Lin [1971] realizan varias propuestas para resolver este problema. A continuación se realiza una extensión de una de estas propuestas para el caso multivariado, con la cual se pueden obtener estimaciones de los parámetros a_{11}, \dots, a_{HH} .

El método propuesto utiliza un algoritmo iterativo que parte de la ecuación (68) en forma agregada:

$$\mathcal{C}\mathbf{y} = \mathcal{C}\mathbf{w}\boldsymbol{\beta} + \mathcal{C}\mathbf{u} \quad (88)$$

Es decir,

$$\mathbf{Y} = \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\mu} \quad (89)$$

La ecuación (89) se puede re-escribir en términos de bloques para cada variable $h, h = 1, \dots, H$ de la siguiente forma:

$$\underbrace{\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_H \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_H \end{bmatrix}}_W \underbrace{\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \\ \vdots \\ \boldsymbol{\beta}_H \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_H \end{bmatrix}}_{\boldsymbol{\mu}} \quad (90)$$

- i. Como primer paso, se estima a través de MCO el modelo (90) para el bloque que corresponde a la variable h para $h = 1, \dots, H$, es decir, $\mathbf{Y}_h = W_h \boldsymbol{\beta}_h + \boldsymbol{\mu}_h$. Este método de estimación supone que $E(\boldsymbol{\mu}_h \boldsymbol{\mu}_h') = V_h^* = \sigma_h^{2*} I_S$. Posteriormente, se generan los residuos $\hat{\boldsymbol{\mu}}_h = [\hat{\mu}_{h,1}, \dots, \hat{\mu}_{h,S}]'$, y se calcula la correlación de primer orden, es decir, $\hat{\rho}_{h,1} = \text{corr}(\hat{\mu}_{h,t}, \hat{\mu}_{h,t-1})$.

Teniendo en cuenta el supuesto sobre los errores de alta frecuencia, $\boldsymbol{\mu}_h = C\mathbf{u}_h$, la estimación de la correlación de primer orden de $\boldsymbol{\mu}_h$ corresponde al elemento de la primera fila y segunda columna de la matriz $E(\boldsymbol{\mu}_h \boldsymbol{\mu}_h') = CV_{hh}C'$, donde V_{hh} es el bloque hh de la matriz V . De esta forma, para el caso de distribución de series trimestrales a mensuales, se tiene la siguiente relación:

$$\hat{\rho}_{h,1} = \frac{a_{hh}^5 + 2a_{hh}^4 + 3a_{hh}^3 + 2a_{hh}^2 + a_{hh}}{2a_{hh}^2 + 4a_{hh} + 3} \quad (91)$$

Por lo tanto,

$$a_{hh}^5 + 2a_{hh}^4 + 3a_{hh}^3 + 2a_{hh}^2(1 - \hat{\rho}_{h,1}) + a_{hh}(1 - 4\hat{\rho}_{h,1}) - 3\hat{\rho}_{h,1} = 0 \quad (92)$$

Un estimador de a_{hh} , para $h = 1, \dots, H$, se puede encontrar resolviendo para a_{hh} la ecuación (92).¹⁰

Por otro lado, para el caso de distribución de series trimestrales a mensuales, los elementos de la matriz Σ^ε se pueden estimar de la siguiente forma:

$$\hat{\sigma}_{hk}^\varepsilon = \hat{\sigma}_{hk}^*/3 \quad (93)$$

donde $\hat{\sigma}_{hk}^* = \frac{1}{S} (\hat{\boldsymbol{\mu}}_h - \hat{a}_{hh}\hat{\boldsymbol{\mu}}_{h,-1})' (\hat{\boldsymbol{\mu}}_k - \hat{a}_{kk}\hat{\boldsymbol{\mu}}_{k,-1})$, $\hat{\boldsymbol{\mu}}_{h,-1} = [\hat{\mu}_{h,0}, \dots, \hat{\mu}_{h,S-1}]'$ y $h, k = 1, \dots, H$.¹¹

Una vez estimados los coeficientes de las matrices A_1 y Σ^ε se pueden obtener estimaciones iniciales del vector de parámetros asociados a las series indicadoras, $\hat{\boldsymbol{\beta}}^{(0)}$, y del vector de series desagregadas, $\hat{\mathbf{y}}^{(0)}$, utilizando las ecuaciones (74) y (73), respectivamente.

¹⁰Chow y Lin [1971] muestran que la solución factible de este polinomio es única.

¹¹El resultado de la ecuación (93) se obtiene debido a que $E(\boldsymbol{\mu}_h - a_{hh}\boldsymbol{\mu}_{h,-1})'(\boldsymbol{\mu}_k - a_{kk}\boldsymbol{\mu}_{k,-1}) = CE(\mathbf{u}_h - a_{hh}\mathbf{u}_{h,-1})'(\mathbf{u}_k - a_{kk}\mathbf{u}_{k,-1})C'$. Por lo tanto, $\sigma_{hk}^* I = \sigma_{hk}^\varepsilon CC'$. Finalmente, para el ejercicio de distribución de series trimestrales a mensuales se tiene que $CC' = 3 I$.

- ii. Posteriormente, se generan unos nuevos residuos del modelo (89) a partir de la estimación inicial del vector de parámetros asociado a las series indicadoras:

$$\hat{\boldsymbol{\mu}}_h^{(1)} = \mathbf{Y}_h - W_h \hat{\boldsymbol{\beta}}_h^{(0)}, \quad h = 1 \dots, H \quad (94)$$

Al igual que la etapa anterior se estiman los coeficientes del vector $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}^{(1)}$, y el vector de series desagregadas, $\hat{\mathbf{y}}^{(1)}$.

- iii. Se repite el paso anterior R veces, hasta obtener convergencia en el vector $\hat{\boldsymbol{\beta}}^{(R)}$ y en el vector de series desagregadas, $\hat{\mathbf{y}}^{(R)}$.

4.4. Método de Moauro y Savio [2005]. Estos autores proponen una metodología de desagregación temporal multivariada basada en modelos de series de tiempo aparentemente no relacionadas o SUTSE. Los autores utilizan un modelo de tendencia lineal local multivariado sobre un vector de series de tiempo $\mathbf{y}_t = [y_{1,t}, \dots, y_{H,t}]'$. Este modelo es especificado de la siguiente forma:

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\xi}_t \quad (95)$$

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_t + \boldsymbol{\eta}_t \quad (96)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t \quad (97)$$

donde $\boldsymbol{\beta}_t$ es la pendiente estocástica, $\boldsymbol{\xi}_t \sim NID(\mathbf{0}, \Sigma_\xi)$, $\boldsymbol{\eta}_t \sim NID(\mathbf{0}, \Sigma_\eta)$ y, $\boldsymbol{\zeta}_t \sim NID(\mathbf{0}, \Sigma_\zeta)$, son vectores aleatorios de dimensión $H \times 1$ no correlacionados entre si con matriz de varianzas y covarianzas Σ_j para $j = \xi, \eta, \zeta$, respectivamente. El modelo puede escribirse en forma estado-espacio como sigue:

$$\mathbf{y}_t = Z_t \boldsymbol{\alpha}_t + G_t \boldsymbol{\varepsilon}_t \quad (98)$$

$$\boldsymbol{\alpha}_{t+1} = T_t \boldsymbol{\alpha}_t + F_t \boldsymbol{\varepsilon}_t \quad (99)$$

donde el vector de estado está definido como: $\boldsymbol{\alpha}_t = [\boldsymbol{\mu}_t', \boldsymbol{\beta}_t']'$, $\boldsymbol{\alpha}_1 \sim NID(\mathbf{0}, P)$, y se tiene que $\boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, I_H)$. Las matrices del sistema están definidas como:

$$T_t = \begin{bmatrix} I_H & I_H \\ 0 & I_H \end{bmatrix}, \quad F_t = \begin{bmatrix} \Gamma_\eta & \Gamma_\zeta \\ 0 & \Gamma_\zeta \end{bmatrix}, \\ Z_t = [I_H \quad 0], \quad G_t = [0 \quad 0 \quad \Gamma_\xi]$$

donde las matrices Γ_j son triangulares inferiores tales que $\Sigma_j = \Gamma_j \Gamma_j'$ para $j = \xi, \eta, \zeta$.

Para el caso de interpolación en el marco de desagregación temporal, el modelo definido en (98)-(99) es válido, sin embargo para el caso de distribución se deben realizar algunas modificaciones.

Para plantear el problema de agregación se define δ como la frecuencia de la serie desagregada ($y_{i,t}$), y S_i^+ como la frecuencia de la serie agregada ($Y_{i,t}$). Las frecuencias deben satisfacer que el

cociente $M_i = \delta/S_i^+$ sea un número entero para $i = 1, \dots, H$.¹² De esta forma, la serie agregada está dada por:

$$Y_{i,t} = \sum_{r=0}^{M_i-1} y_{i,t-r} \quad t = M_i, 2M_i, \dots \quad i = 1, \dots, H \quad (100)$$

El proceso generador de las series Y_t está dado por las ecuaciones (95) a (97), sin embargo algunos de sus componentes se observan de forma agregada temporalmente, lo que implica que el vector \mathbf{y}_t contiene series observadas en alta y en baja frecuencia. En este caso el acumulador de la serie $y_{i,t}$ se define como:

$$y_{i,t-M_i+r}^c = \sum_{j=1}^r y_{i,t-M_i+j} \quad r = 1, \dots, M_i \quad i = 1, \dots, H \quad (101)$$

lo que implica que $y_{i,t}^c = Y_{i,t}$ para $t = M_i, 2M_i, \dots$. En forma matricial, el acumulador se puede escribir como:

$$\mathbf{y}_t^c = C_t \mathbf{y}_{t-1}^c + \boldsymbol{\mu}_t + \boldsymbol{\xi}_t \quad (102)$$

$$= C_t \mathbf{y}_{t-1}^c + \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t + \boldsymbol{\zeta}_t + \boldsymbol{\xi}_t \quad (103)$$

donde C_t es una matriz diagonal. Los elementos de la diagonal de esta matriz están dados por el vector $[c_{1,t}, \dots, c_{H,t}]'$ y cumplen la siguiente condición para $i = 1, \dots, H$:

$$c_{i,t} = \begin{cases} 0 & \text{si } t = 1, M_i + 1, 2M_i + 1, \dots \\ 1 & \text{si en otro caso} \end{cases} \quad (104)$$

Para incluir a $\mathbf{y}_{i,t}$ en el modelo definido en (98)-(99), se deben redefinir las matrices del modelo de estado espacio de la siguiente forma:

$$\boldsymbol{\alpha}_t = \begin{bmatrix} \boldsymbol{\mu}_t \\ \boldsymbol{\beta}_t \\ \mathbf{y}_t^c \end{bmatrix}, T_t = \begin{bmatrix} I_H & I_H & 0 \\ 0 & I_H & 0 \\ I_H & I_H & C_t \end{bmatrix}, F = \begin{bmatrix} \Gamma_\eta & 0 & 0 \\ 0 & \Gamma_\zeta & 0 \\ \Gamma_\eta & \Gamma_\zeta & \Gamma_\xi \end{bmatrix}, Z = [0 \ 0 \ I_H], G = 0 \quad (105)$$

con $\mathbf{y}_0^c = \mathbf{0}$. Es importante notar que en esta versión del modelo, la matriz T es variable en el tiempo. Finalmente, la estimación de la serie desagregada se obtiene por máxima verosimilitud utilizando el filtro de Kalman.

Para mayor claridad de la construcción de las matrices anteriormente definidas, suponga que se quiere efectuar la desagregación de dos series ($y_{1,t}, y_{2,t}$) trimestrales a mensuales usando dos series indicadoras ($y_{3,t}, y_{4,t}$) en frecuencia mensual. En este caso, se tiene que: $H = 4$, $\delta = 12$, $S_1^+ = S_2^+ = 4$, $S_3^+ = S_4^+ = 12$, $M_1 = M_2 = 3$, $M_3 = M_4 = 1$. Además, con $\mathbf{c}_i = [c_{i,1}, c_{i,2}, c_{i,3}, \dots, c_{i,T}]'$ para $i = 1, 2, 3, 4$ se tiene que $\mathbf{c}_1 = \mathbf{c}_2 = [0, 1, 1, 0, 1, 1, 0, 1, 1, \dots]$ y $\mathbf{c}_3 = \mathbf{c}_4 = [0, 0, 0, 0, 0, 0, 0, 0, \dots]$.

¹²Por ejemplo, para $i = 1$ si la frecuencia de la serie desagregada es mensual, $\delta = 12$ y la frecuencia de la serie agregada es trimestral, $S_1^+ = 4$, entonces el cociente de frecuencias es $M_1 = 12/4 = 3$.

Como lo señalan los autores, una de las ventajas del modelo propuesto es que no impone una estructura *a priori* en los datos, ya que la especificación del modelo tiene en cuenta las características estocásticas de las series. Adicionalmente, las metodologías de Chow y Lin [1971], Fernández [1981] y Litterman [1983] se pueden expresar en términos de los modelos SUTSE.

Por otro lado, una de las desventajas de este método es que requiere especificar los valores iniciales del vector de estado y de la matriz de error cuadrático medio asociada.

5. ILUSTRACIÓN EMPÍRICA

En esta sección se aplica el método multivariado de Di-Fonzo [1990] y la extensión propuesta en este documento a las cuentas nacionales colombianas. La aplicación consiste en estimar las series mensuales asociadas al Producto Interno Bruto (PIB), consumo de los hogares, consumo del gobierno y formación bruta de capital a partir de las series trimestrales correspondientes, así como cinco series mensuales, cuatro variables indicadoras y las exportaciones netas. Debido a que las series a desagregar se definen como variables flujo, el problema planteado en esta sección corresponde a un problema de distribución.

Las series indicadoras (mensuales) son el Índice de Producción Real de la industria manufacturera colombiana sin trilla de café (IPR), el índice del comercio minorista sin combustibles ni vehículos, los pagos totales efectuados por el gobierno sin inversión, y las importaciones de bienes de capital. Estas cuatro variables están asociadas al PIB, al consumo de los hogares, al consumo del gobierno y a la formación bruta de capital, respectivamente.

Para aplicar los métodos de desagregación temporal multivariados descritos en la sección anterior, se debe contar con una restricción contemporánea. Para ello, se debe tener en cuenta que en una economía abierta con gobierno existe la siguiente relación:

$$Y + M = C_p + G + I_p + X \quad (106)$$

$$Y = C_p + G + I_p + X - M \quad (107)$$

La ecuación (106) dice que la suma del producto total de la economía (Y) y las importaciones (M), se destina al consumo privado (C_p), al gasto del gobierno (G), a la inversión privada (I_p) y a las exportaciones (X). Además, teniendo en cuenta que el gasto del gobierno se puede desagregar en gasto público de inversión (I_g) y gasto público de consumo (C_g), se tiene que $G = C_g + I_g$. Si se define la inversión total (I) como la suma de la inversión pública más la inversión privada, y las exportaciones netas (XN) como las exportaciones menos las importaciones, entonces la ecuación (107) se puede escribir como:

$$Y = C_p + C_g + I_p + I_g + X - M \quad (108)$$

$$Y = C_p + C_g + I + XN \quad (109)$$

$$XN = Y - C_p - C_g - I \quad (110)$$

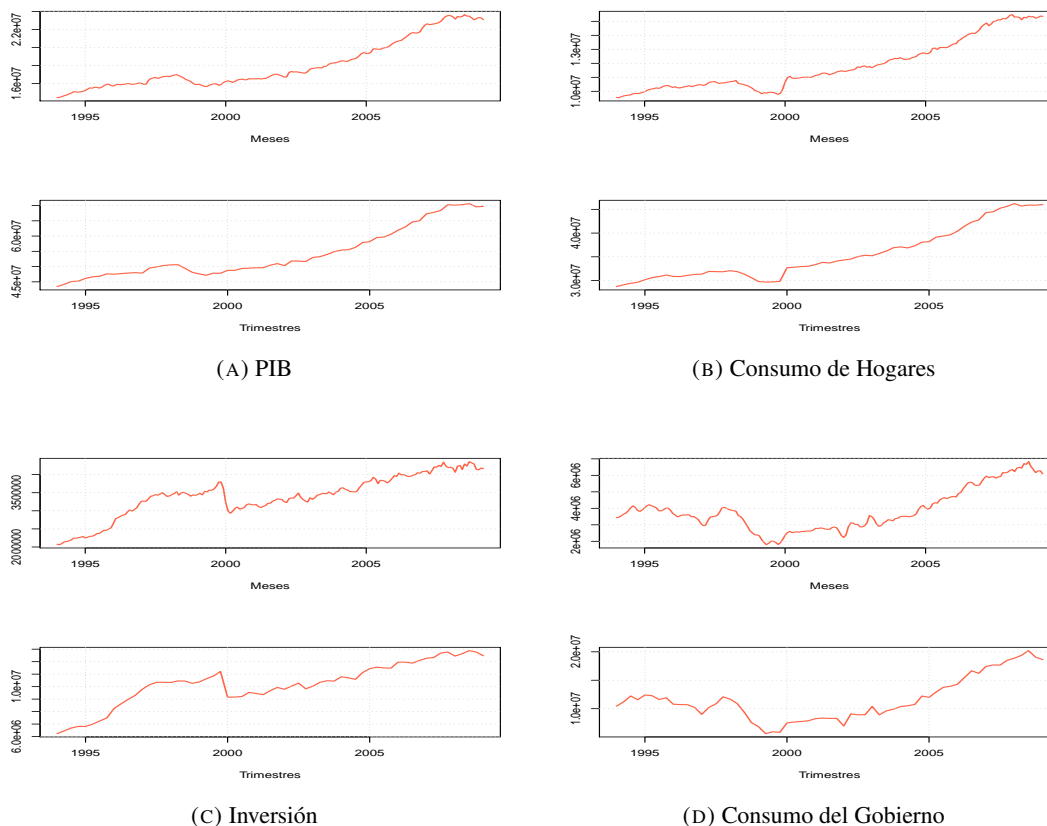
De esta forma, en términos reales para las cuentas nacionales, se debe cumplir que el PIB menos el consumo de los hogares, menos el consumo del gobierno, menos la formación bruta de capital sea igual a las exportaciones netas para todos los periodos del tiempo. Por lo tanto, el agregado contemporáneo de esta aplicación es la serie de exportaciones netas reales en frecuencia mensual, que corresponde a la variable z definida en (69).

La muestra trimestral incluye el periodo comprendido entre el primer trimestre de 1994 y el primer trimestre de 2009. Por otro lado, la muestra mensual contiene información desde enero de 1994 hasta marzo de 2009. Los datos trimestrales estaban disponibles de forma desestacionalizada mientras que los datos mensuales fueron ajustados estacionalmente utilizando el procedimiento X12. Adicionalmente, todas las series se tomaron en precios constantes del año 2000.¹³

Los resultados muestran que las series estimadas de alta frecuencia siguen la misma tendencia que sus contrapartes de baja frecuencia, pero con mayor volatilidad. Para disminuir la volatilidad, se aplicó el filtro de Hodrick-Prescott a las series indicadoras. Los resultados de las series suavizadas muestran una menor volatilidad, pero siguen la misma tendencia que las series no suavizadas. La Gráfica 1 muestra las series estimadas de alta frecuencia para el PIB, el consumo de los hogares, la inversión y el consumo del gobierno. Es de notar que la metodología utilizada supone que los errores de alta frecuencia siguen un proceso $VAR(1)$.

Por otro lado, si se supone que los errores de alta frecuencia siguen un proceso ruido blanco multivariado, se obtienen estimaciones de las series desagregadas que presentan una mayor volatilidad que bajo el supuesto de $VAR(1)$. Por ejemplo, en la Gráfica 2 se comparan las series de alta frecuencia estimadas para el PIB con cada uno de los supuestos sobre el proceso del error, y se observa claramente una mayor volatilidad en la serie estimada suponiendo ruido blanco. Cabe anotar que este resultado también se obtiene si no se realiza el suavizamiento de las variables indicadoras por medio del filtro de Hodrick y Prescott.

¹³Debido a diferencias metodológicas la serie de exportaciones netas en frecuencia trimestral no es exactamente igual a la serie agregada mensual. Por lo tanto, se ajustó la serie de exportaciones netas en frecuencia mensual con el fin de que su agregación coincidiera con la serie en frecuencia trimestral. El ajuste se realizó utilizando el método de Chow y Lin [1971].



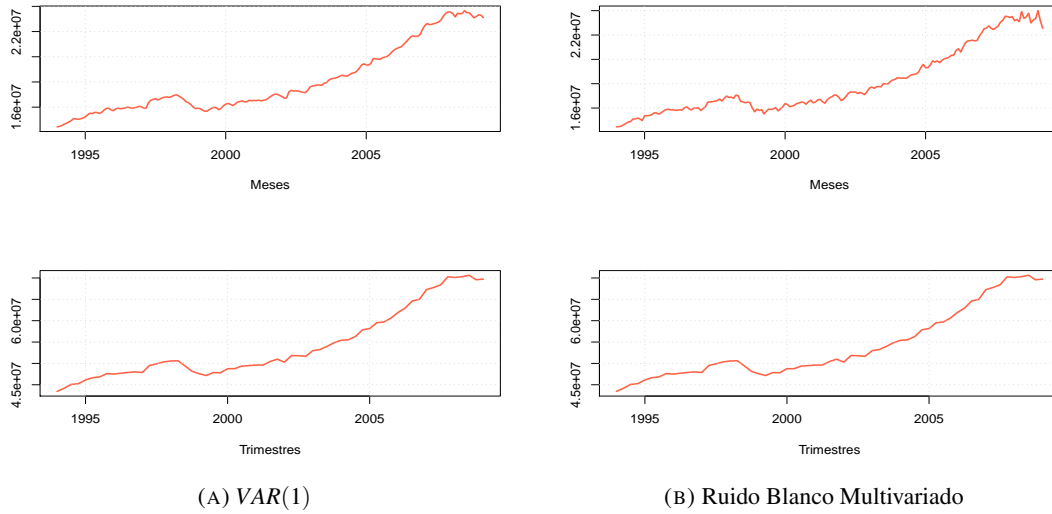
GRÁFICA 1. Series trimestrales observadas y series mensuales estimadas suponiendo que los errores de alta frecuencia siguen un proceso $VAR(1)$.

6. COMENTARIOS FINALES

Los resultados de este documento resaltan la importancia de estudiar las características de las series a desagregar, con el objeto de aplicar el método más apropiado. Cada uno de estos métodos implica supuestos diferentes sobre las series de baja frecuencia e indicadores. Por esto, es importante tener en cuenta características estocásticas de las series tales como el orden de integración y posibles relaciones de cointegración.

Adicionalmente, es recomendable utilizar metodologías multivariadas y no univariadas puesto que las primeras permiten modelar dinámicas más generales. Por otro lado, las metodologías multivariadas pueden tener en consideración restricciones contemporáneas sobre las series.

En este documento se propone una extensión de la metodología de desagregación multivariada de Di-Fonzo [1990]. Esta supone que los errores de las series de alta frecuencia siguen un modelo



GRÁFICA 2. Series estimadas del PIB suponiendo que los errores de alta frecuencia siguen un proceso ruido blanco multivariado y un VAR(1).

VAR(1) en lugar de un proceso ruido blanco. Finalmente, la metodología propuesta es aplicada a las cuentas nacionales colombianas trimestrales con el objeto de generar estimaciones de los datos mensuales.

REFERENCIAS

- AADLAND, D. (2000): "Distribution and Interpolation Using Transformed Data," *Journal of Applied Statistics*, 27, 141–56.
- BARCELLAN, R. (2002): "ECOTRIM: a program for temporal disaggregation of time series," in *Proceedings of the Workshop on Quarterly National Accounts, Luxembourg: Eurostat*, ed. by R. Barcellan, y G. Mazzi, pp. 79–95.
- CHOW, G., y A. LIN (1971): "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series," *Review of Economics and Statistics*, 53(4), 372–375.
- DENTON, F. (1971): "Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization," *Journal of the American Statistical Association*, 66(333), 99–102.
- DI-FONZO, T. (1987): *La stima indiretta di serie economiche trimestrali*. CLEUP.
- (1990): "The estimation of M disaggregate time series when contemporaneous and temporal aggregates are known," *The Review of Economics and Statistics*, 72(1), 178–182.
- (2002): "Temporal disaggregation of a system of time series when the aggregate is known: optimal vs. adjustment methods," in *Proceedings of the Workshop on Quarterly National Accounts, Eurostat*, ed. by R. Barcellan, y G. Mazzi, pp. 63–77.
- FERNÁNDEZ, R. (1981): "A methodological note on the estimation of time series," *The Review of Economics and Statistics*, 63(3), 471–476.
- GUERRERO, V. (1990): "Temporal disaggregation of time series: an ARIMA-based approach," *International Statistical Review*, 58, 111–120.
- LITTERMAN, R. (1983): "A random walk, Markov model for the distribution of time series," *Journal of Business and Economic Statistics*, 1(2), 169–173.
- LUPI, C., y G. PARIGI (2002): "Temporal Disaggregation of Economic Time Series: Some Econometric Issues," in *Proceedings of the Workshop on Quarterly National Accounts, Luxembourg: Eurostat*, ed. by R. Barcellan, y G. Mazzi, pp. 111–140.
- MOAURO, F., y G. SAVIO (2005): "Temporal disaggregation using multivariate structural time series models," *Econometrics Journal*, 8(2), 214–234.
- PROIETTI, T. (1999): "Distribution and interpolation revisited: a structural approach," *Statistica*, 58, 411–432.
- ROSSI, N. (1982): "A note on the estimation of disaggregate time series when the aggregate is known," *The Review of Economics and Statistics*, 64(4), 695–696.
- SANTOS-SILVA, J., y F. CARDOSO (2001): "The Chow-Lin method using dynamic models," *Economic Modelling*, 18, 269–280.

APÉNDICE A. MATRIZ DE COVARIANZAS DE LOS ERRORES DEL MODELO PROPUESTO

En este apéndice se deduce la estructura de la matriz de varianzas y covarianzas de los errores del modelo propuesto en la sección 4.3, en el cual se supone que éstos siguen un modelo VAR(1).

Reescribiendo la ecuación (81) en notación matricial para todos los periodos de tiempo se obtiene:

$$\mathbf{u}^\bullet = (I_N \otimes A_1) \mathbf{u}_{-1}^\bullet + \boldsymbol{\varepsilon}^\bullet \quad (111)$$

donde $\mathbf{u}^\bullet = [\mathbf{u}_1^\bullet, \mathbf{u}_2^\bullet, \dots, \mathbf{u}_N^\bullet]'$, $\mathbf{u}_{-1}^\bullet = [\mathbf{u}_0^\bullet, \mathbf{u}_1^\bullet, \dots, \mathbf{u}_{N-1}^\bullet]'$ y $\boldsymbol{\varepsilon}^\bullet = [\boldsymbol{\varepsilon}_1^\bullet, \boldsymbol{\varepsilon}_2^\bullet, \dots, \boldsymbol{\varepsilon}_N^\bullet]'$.

Debido a que el proceso es estacionario, la matriz de varianzas y covarianzas de $\boldsymbol{\varepsilon}^\bullet$ se puede definir como sigue:

$$V^\bullet = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{N-1} \\ \Gamma_1' & \Gamma_0 & \Gamma_1 & \cdots & \Gamma_{N-2} \\ \Gamma_2' & \Gamma_1' & \Gamma_0 & \cdots & \Gamma_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Gamma_{N-1}' & \Gamma_{N-2}' & \Gamma_{N-3}' & \cdots & \Gamma_0 \end{bmatrix} \quad (112)$$

donde $\Gamma_i = E(\mathbf{u}_t^\bullet \mathbf{u}_{t-i}^{\bullet'})$, para $i = 0, \dots, N-1$, es una matriz de dimensión $H \times H$ definida como:

$$\Gamma_i = \begin{bmatrix} E(u_{1,t} u_{1,t-i}) & E(u_{1,t} u_{2,t-i}) & \cdots & E(u_{1,t} u_{H,t-i}) \\ E(u_{2,t} u_{1,t-i}) & E(u_{2,t} u_{2,t-i}) & \cdots & E(u_{2,t} u_{H,t-i}) \\ \vdots & \vdots & \ddots & \vdots \\ E(u_{H,t} u_{1,t-i}) & E(u_{H,t} u_{2,t-i}) & \cdots & E(u_{H,t} u_{H,t-i}) \end{bmatrix} = \begin{bmatrix} \gamma_{11}^{(i)} & \gamma_{12}^{(i)} & \cdots & \gamma_{1H}^{(i)} \\ \gamma_{21}^{(i)} & \gamma_{22}^{(i)} & \cdots & \gamma_{2H}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{H1}^{(i)} & \gamma_{H2}^{(i)} & \cdots & \gamma_{HH}^{(i)} \end{bmatrix}$$

A partir de los anteriores resultados, se puede mostrar que la matriz de varianzas y covarianzas del vector \mathbf{u} definido en (68) es de la forma:

$$V = \begin{bmatrix}
 \gamma_{11}^{(0)} & \gamma_{11}^{(1)} & \cdots & \gamma_{11}^{(N-1)} & \gamma_{12}^{(0)} & \gamma_{12}^{(1)} & \cdots & \gamma_{12}^{(N-1)} & \cdots & \gamma_{1H}^{(0)} & \gamma_{1H}^{(1)} & \cdots & \gamma_{1H}^{(N-1)} \\
 \gamma_{11}^{(1)} & \gamma_{11}^{(0)} & \cdots & \gamma_{11}^{(N-2)} & \gamma_{21}^{(1)} & \gamma_{12}^{(0)} & \cdots & \gamma_{12}^{(N-2)} & \cdots & \gamma_{H1}^{(1)} & \gamma_{1H}^{(0)} & \cdots & \gamma_{1H}^{(N-2)} \\
 \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
 \gamma_{11}^{(N-1)} & \gamma_{11}^{(N-2)} & \cdots & \gamma_{11}^{(0)} & \gamma_{21}^{(N-1)} & \gamma_{21}^{(N-2)} & \cdots & \gamma_{12}^{(0)} & \cdots & \gamma_{H1}^{(N-1)} & \gamma_{H1}^{(N-2)} & \cdots & \gamma_{1H}^{(0)} \\
 \gamma_{21}^{(0)} & \gamma_{21}^{(1)} & \cdots & \gamma_{21}^{(N-1)} & \gamma_{22}^{(0)} & \gamma_{22}^{(1)} & \cdots & \gamma_{22}^{(N-1)} & \cdots & \gamma_{2H}^{(0)} & \gamma_{2H}^{(1)} & \cdots & \gamma_{2H}^{(N-1)} \\
 \gamma_{12}^{(1)} & \gamma_{21}^{(0)} & \cdots & \gamma_{21}^{(N-2)} & \gamma_{22}^{(1)} & \gamma_{22}^{(0)} & \cdots & \gamma_{22}^{(N-2)} & \cdots & \gamma_{H2}^{(1)} & \gamma_{2H}^{(0)} & \cdots & \gamma_{2H}^{(N-2)} \\
 \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
 \gamma_{12}^{(N-1)} & \gamma_{12}^{(N-2)} & \cdots & \gamma_{21}^{(0)} & \gamma_{22}^{(N-1)} & \gamma_{22}^{(N-2)} & \cdots & \gamma_{22}^{(0)} & \cdots & \gamma_{H2}^{(N-1)} & \gamma_{H2}^{(N-2)} & \cdots & \gamma_{2H}^{(0)} \\
 \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
 \gamma_{H1}^{(0)} & \gamma_{H1}^{(1)} & \cdots & \gamma_{H1}^{(N-1)} & \gamma_{H2}^{(0)} & \gamma_{H2}^{(1)} & \cdots & \gamma_{H2}^{(N-1)} & \cdots & \gamma_{HH}^{(0)} & \gamma_{HH}^{(1)} & \cdots & \gamma_{HH}^{(N-1)} \\
 \gamma_{1H}^{(1)} & \gamma_{H1}^{(0)} & \cdots & \gamma_{H1}^{(N-2)} & \gamma_{2H}^{(1)} & \gamma_{H2}^{(0)} & \cdots & \gamma_{H2}^{(N-2)} & \cdots & \gamma_{HH}^{(1)} & \gamma_{HH}^{(0)} & \cdots & \gamma_{HH}^{(N-2)} \\
 \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
 \gamma_{1H}^{(N-1)} & \gamma_{1H}^{(N-2)} & \cdots & \gamma_{H1}^{(0)} & \gamma_{2H}^{(N-1)} & \gamma_{2H}^{(N-2)} & \cdots & \gamma_{H2}^{(0)} & \cdots & \gamma_{HH}^{(N-1)} & \gamma_{HH}^{(N-2)} & \cdots & \gamma_{HH}^{(0)}
 \end{bmatrix} \quad (113)$$

Al considerar un modelo VAR(1) se tiene que $\Gamma_i = A_1^i \Gamma_0$ para $i \in \mathbb{N}$. Adicionalmente, si el modelo VAR(1) es diagonal se obtiene que $\gamma_{hk}^{(i)} = a_{hh}^i \gamma_{hk}^{(0)}$, para $i = 1, 2, \dots$ y $k, h = 1, \dots, H$. Por lo tanto, para este tipo de modelos se tiene que la matriz de covarianzas V depende de los elementos de las matrices Γ^0 y A_1 de la siguiente forma:

$$V = \begin{bmatrix}
 \gamma_{11}^{(0)} A_1^{(11)} & \gamma_{12}^{(0)} A_1^{(01)} + \gamma_{21}^{(0)} A_1^{(20)} & \cdots & \gamma_{1H}^{(0)} A_1^{(01)} + \gamma_{H1}^{(0)} A_1^{(H0)} \\
 \gamma_{21}^{(0)} A_1^{(02)} + \gamma_{12}^{(0)} A_1^{(10)} & \gamma_{22}^{(0)} A_1^{(22)} & \cdots & \gamma_{2H}^{(0)} A_1^{(02)} + \gamma_{H2}^{(0)} A_1^{(H0)} \\
 \vdots & \vdots & \ddots & \vdots \\
 \gamma_{H1}^{(0)} A_1^{(0H)} + \gamma_{1H}^{(0)} A_1^{(10)} & \gamma_{H2}^{(0)} A_1^{(0H)} + \gamma_{2H}^{(0)} A_1^{(20)} & \cdots & \gamma_{HH}^{(0)} A_1^{(HH)}
 \end{bmatrix} \quad (114)$$

donde:

$$A_1^{(hh)} = \begin{bmatrix}
 1 & a_{hh} & a_{hh}^2 & \cdots & a_{hh}^{N-1} \\
 a_{hh} & 1 & a_{hh} & \cdots & a_{hh}^{N-2} \\
 a_{hh}^2 & a_{hh} & 1 & \cdots & a_{hh}^{N-3} \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 a_{hh}^{N-1} & a_{hh}^{N-2} & a_{hh}^{N-3} & \cdots & 1
 \end{bmatrix} \quad (115)$$

$$A_1^{(0h)} = \begin{bmatrix} 1 & a_{hh} & a_{hh}^2 & \cdots & a_{hh}^{N-1} \\ 0 & 1 & a_{hh} & \cdots & a_{hh}^{N-2} \\ 0 & 0 & 1 & \cdots & a_{hh}^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad (116)$$

$$A_1^{(h0)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{hh} & 0 & 0 & \cdots & 0 \\ a_{hh}^2 & a_{hh} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{hh}^{N-1} & a_{hh}^{N-2} & a_{hh}^{N-3} & \cdots & 0 \end{bmatrix} \quad (117)$$

para $h = 1, \dots, H$.