MODELOS ESTOCASTICOS DE CAJA NEGRA

JORGE ARAVENA L.
DEPARTAMENTO DE INGENIERIA ELECTRICA
UNIVERSIDAD DE SANTIAGO, CHILE

RESUMEN

En este trabajo se analiza el problema de determinar modelos matemáticos, realizables, que permitan relacionar dos procesos estocásticos. De estos procesos solo se requiere conocer las estadísticas de segundo orden de modo que la metodología es muy adecuada para el trabajo experimental como el que se usa en métodos de Caja Negra.

Se hace una formulación empleando los conceptos de espacios de resolución y teoría de causalidad. Este enfoque permite una gran generalidad en el planteamiento y de hecho extiende resultados previamente conocidos sólo para procesos estacionarios. Al mismo tiempo permite imponer la condición de causalidad en el planteamiento mismo del problema asegurando la realizabilidad de los modelos.

La metodología seguida pone en evidencia los distintos tipos de modelos que se pueden plantear, como se relacionan entre si y las suposiciones adicionales que es necesario realizar en cada caso.

1.- INTRODUCCION.

El análisis de sistemas desde un punto de vista generalizado permite abstraer aquellas propiedades que son escenciales en un problema dado. Es posible en muchos casos establecer un nexo entre problemas de muy distintas disciplinas. El caso que aquí se estudiará es común, por ejemplo, en áreas de economía y bioingeniería [1] , [2] , [3]. Se resolverá empleando un enfoque en términos de operadores en espacios de funciones.

Para describir adecuadamente la evolución dinámica en el tiempo se empleará como marco matemático a los espacios Hilbert de resolución. De esta manera se podrá incluir en la formulación la restricción de causalidad de los modelos. En la sección 2 se presentan en forma muy sintética los conceptos matemáticos principales.

El problema que se resolverá en este trabajo se puede plantear informalmente de la siguiente manera: Dados dos procesos estocásticos (u) e(y), determinar un modelo matemático fisicamente realizable que los relacione.

El origen de este problema y su motivación principal es en la modelación de sistemas sobre los cuales se tiene una muy limitada descripción fenomeno lógica. En casos extremos puede existir duda sobre la relación causa/efecto entre las variable.

Por ejemplo en [1] y [3] se reportan análisis de la relación entre el ingreso promedio nacional y la oferta de dinero en el mercado de capitales. En el caso de [1] se concluyó que en Estados Unidos no hay efecto del ingreso sobre la oferta de dinero. En el caso de Gran Bretaña [3] se obtiene un resultado diferente. En ambos casos se detecta efecto de la oferta de dinero sobre el ingreso nacional. Se puede decir que en un caso no hay realimentación entre las variables pero si la hay en el segundo caso.

La detección de realimentación entre variables ha sido el problema que ha concentrado la atención en esta línea ([4], [5], [6]). La existencia de realimentaciones puede ser decisiva en la elección de esquemas o políticas de control.

Hasta el momento se ha considerado solamente el caso de procesos estocásticos estacionarios requiriendo condiciones de triangularidad en matrices de funciones de transferencia ([4], [5]) ó en factorizaciones de la densidad espectral conjunta ([6]).

El presente trabajo también emplea factorización. Sin embargo ésta se aplica directamente a los operadores de covariancia e incluye restricciones de causalidad para asegurar que los modelos son realizables. Los resultados que se obtienen son válidos tanto para procesos estacionarios como no estacionarios y por lo tanto son una extensión sobre los resultados publicados en la literatura.

De particular novedad son los resultados que relacionan modelos de lazo abierto con el concepto existente de ausencia de realimentación en sentido débil. También, al emplear la formulación en base a operadores causales, la relación entre el problema de ausencia de realimentación y el problema de Wiener de estimación estadística surge en forma natural.

Dado que se emplean estadísticas de segundo orden solamente, la metodología es aplicable con relativa facilidad cuando se poseen registros temporales (series) de las variables que se desea analizar. Por la misma razón existirán situaciones que no son discriminables con la información usada. Este problema es ilustrado con un ejemplo.

2.- ANTECEDENTES MATEMATICOS.

Se presentan aquí los conceptos más básicos referentes a espacios de resolución. Por razones de espacio esta presentación es necesariamente axiomática. Mayores antecedentes se pueden encontrar por ejemplo en [8] y [10] .

Espacio de resolución discreta es un par (H, Δ) en que H es un espacio de Hilbert (separable) y $\Delta = \{\Delta(k), k = 0,1,\ldots\}$ es un conjunto de ortoproyectores en H y que satisface

$$\Delta(i)\,\Delta(j) = 0 \quad i \neq j$$
$$I = \sum_k \Delta(k) \quad .$$

Asociada a la familia Δ se define la familia de ortoproyectores $P = \{P^k\}$ usando

$$P^k = \sum_{i \leq k} \Delta(i).$$
$$= P^{k-1} + \Delta(k).$$

Dado el vector $x \in H$, en forma intuitiva se dice que $x_k = \Delta(k)x$ es el "valor de x en el instante k". Por extensión al vector $P^{k-1}x$ se le denomina el pasado estricto de x. El vector $P^k x$ contiene el pasado estricto más el valor presente de x.

Una transformación T lineal continua en H se define como causal si satisface

$$\Delta(k)T = \Delta(k)TP^k \quad \forall k.$$

A menos que se diga explícitamente lo contrario los términos transformación y operador implicarán siempre linealidad y continuidad.

Básicamente esta definición establece que el valor presente de la salida de T se puede calcular sin conocer el futuro.

Si la transformación T satisface la condición $\Delta(k)T = \Delta(k)\ TP^{k-1}$ se denomina estrictamente causal. En este caso el valor presente de la salida no depende del valor presente de la entrada sino sólo de valores estrictamente en el pasado.

Se define como (estrictamente) anticausal a una transformación T tal que su adjunta T^* es (estrictamente) causal.

Un operador que es simultáneamente causal y anticausal se denomina sin memoria o estático. En este caso el valor presente de la salida depende sólo del valor presente de la entrada.

Descomposición Aditiva de un operador T es su representación en la forma

$$T = T_Z + T_{\underline{C}} + T_{\underline{C}*}$$

En que T_Z es sin memoria, $T_{\underline{C}}$ es estrictamente causal y $T_{\underline{C}*}$ es estrictamente anticausal.

Un operador T tal que T^{-1} es continuo admite una factorización regular si es posible escribir

$$T = A\ B^*$$

En que A, A^{-1}, B, B^{-1} son todos operadores causales.

Una factorización de particular importancia ocurre cuando $T > 0$ y $A = B$. En este caso se tiene $T = AA^*$ y se habla de una factorización espectral del operador T.

El espacio de probabilidad de los procesos estocasticos con valores en H es la terna (H, Σ, p) en que Σ es una familia de conjuntos de Borel y p es una medida de probabilidad

Para cada proceso estocástico $\alpha \in (H, \Sigma, p)$, con momentos de primer y segundo orden finitos, se define el valor medio de α, $m_\alpha \in H$ y el operador de convarianza Q_α mediante

$$E\left[<x,\alpha>\right] = <x,m_\alpha>\ , \quad x \in H$$

$$E\left[<x,\alpha - m_\alpha><y,\alpha - m_\alpha>\right] = <x,Q_\alpha\ y>, \quad x,y \in H.$$

En las expresiones anteriores $<.\ ,\ .>$ indica el producto interno en H y $E[\cdot]$ es el valor esperado de la variable aleatoria indicada.

En forma similar, dados $\alpha,\beta \in (H,\Sigma,p)$ se define la covarianza cruzada $Q_{\alpha\beta}$ mediante

$$E\left[<x,\alpha - m_\alpha><y,\beta - m_\beta>\right] = <x,Q_{\alpha\beta}y>, \quad x,y \in H.$$

Los operadores de varianza son usados fuertemente en este trabajo. Algunas de sus propiedades básicas se listan a continaución. Se supone $\alpha,\beta \in (H,\Sigma,p)$ y L,T operadores en H.

i) $Q_{L_\alpha,\ T_\beta} = L\ Q_{\alpha\beta}\ T^*$

ii) $Q_{\alpha+\beta} = Q_\alpha + Q_{\alpha\beta} + Q_{\beta\alpha} + Q_\beta$

iii) $Q_{\alpha\beta} = Q^*_{\beta\alpha}$

iv) $E\left[<\alpha,\alpha>\right]$ es finito si y sólo si Q_α es operador con traza finita. En este caso se cumple $E\left[<\alpha,\alpha>\right] = tr\ (Q_\alpha)$

Vale la pena recordar también algunas propiedades de la función traza y que serán utilizadas en este trabajo:

a) $tr\ (T) = \Sigma<Te_i,\ e_i>$ para cualquier base ortonormal del espacio H

b) $tr\ (TL) = tr\ (LT)$

c) $tr\ (T+L) = tr\ (T) + tr\ (L)$

3.- MODELOS DE INTERACCIONES.

Para plantear formalmente el problema se supondrán dos procesos estocásticos u e y que toman valores en un espacio de Hilbert H y de los cuales se conocen estadísticas de primero y segundo orden. Por comodidad se supondrán con valor medio nulo de modo que solo se requiere conocer los operadores de covariancia Q_u, Q_y, Q_{yu}.

Si se define $\zeta = (y,u)$ como el proceso conjunto que toma valores en H^2, este proceso tendrá media nula y un operador de covariancia Q_ζ. Empleando notación matricial se tendrá

$$Q_\zeta = \begin{bmatrix} Q_y & Q_{yu} \\ Q_{uy} & Q_u \end{bmatrix} \qquad (1)$$

Se supondrá que este operador es positivo definido ($Q_\zeta > 0$). Esta suposición equivale a la de rango completo que, para el caso estacionario, se hace en [4], [5] y [6].

Es evidente que si H tiene una estructura de resolución discreta definida por los ortoproyectores $\{\Delta(k),\ k = 0, 1, \ldots\}$ se puede inducir en forma muy sencilla una estructura de resolución en H^2. Con un pequeño abuso de notación se tendría:

$$\Delta(k)\zeta = \begin{bmatrix} \Delta(k)y \\ \Delta(k)u \end{bmatrix}$$

Ahora bien, como $Q_\zeta > 0$ (en H^2) entonces admite una factorización espectral [7]. Se puede escribir

$$Q_\zeta = A\, A^* \qquad \text{en } H^2$$

en que A es un operador/causal con inversa causal. Es decir

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Cada una de las entradas A_{ij} debe ser un operador causal en H.

La factorización espectral no es única. Se sabe que si AA^* y BB^* son dos factorizaciones distintas se debe cumplir

$$\begin{aligned} B &= A\,N \\ N\,N^* &= I \\ N &- \text{estático (sin memoria).} \end{aligned}$$

El resultado de factorización es basicamente un procedimiento de "blanqueo" de los procesos estocásticos. Implica que los procesos u e y admiten la representación que se indica en la Figura 1. Este modelo se llamará un modelo directo. Las señales n_1 y n_2 son secuencias de ruido blanco independientes entre si y con covariancia unitaria ($Q_{u_1} = Q_{n_2} = I$, $Q_{n_1 n_2} = 0$). El bloque designado por N corresponde al operador estático que relaciona las distintas factorizaciones.

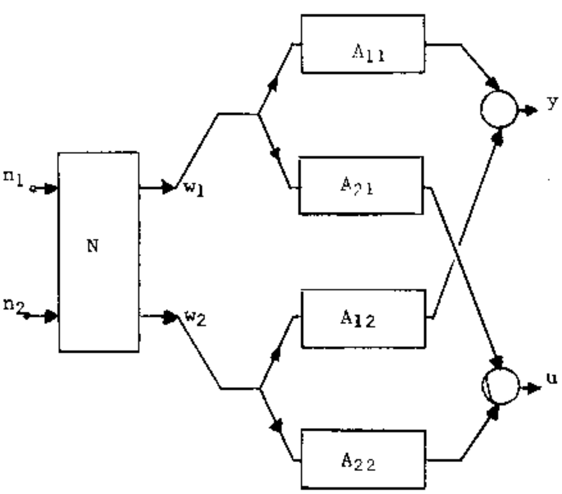

Figura 1: Modelo Directo de Interacción Entre los Procesos u e y.

Desdeel punto de vista de modelación, el esquema de la figura 1 es realizable ya que todos los operadores son causales. La simplementación de este modelo permitirá generar procesos estocásticos que tendrán la misma estadísticas de segundo orden que las dadas como dato (Q_y, Q_u, Q_{yu}).

El modelo directo es especialmente adecuado para trabajos de estimación y predicción del proceso $\zeta = (y,u)$. Cualquiera que sea el bloque N, como $N\,N^* - I$, las señales W_1, W_2 son también ruidos blancos independientes entre si y normalizados.

Desgraciadamente el modelo directo es claramente insuficiente para examinar la relación causa/efecto entre los procesos u e y. Solo permite establecer que ambos procesos tendrían "causas comunes" y que variaciones en uno de ellos deberían ir acompañadas de variaciones en el otro. A pesar de ello es importante destacar que aún con este modelo se ve que la condición $A_{21} = 0$ ($A_{12} = 0$) hace más direccional el modelo. Por ejemplo, toda variación en u necesariamente significará una variación en y pero no a la inversa.

En el modelo directo ambos procesos aparecen como salidas. Si se postula que hay una relación causa/efecto entre ellos es mas razonable tratar de determinar un modelo realimentado como el que se muestra en la Figura 2. En este modelo las señales α y β corresponden a posibles agentes externos.

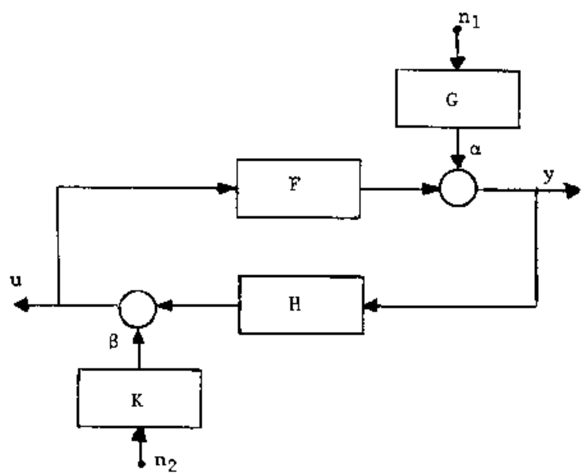

Figura 2: Modelo Realimentado de Interacción.

En este caso surge como importante determinar si por ejemplo se puede suponer H = 0. Se tendría una situación muy clara. El proceso β = u es causa y puede razonablemente ser elegido como variable manipulable para afectar al proceso y. Este punto será examinado en detalle en la siguiente sección.

Es necesario notar que a partir de la información estadística, la técnica de "blanqueo" de los procesos genera siempre un modelo directo. Empleando algrebra de bloques se puede determinar que un modelo directo no siempre se puede transformar en un modelo realimentado. Se requieren suposiciones adicionales especialmente para asegurar la causalidad de los operadores. En particular se tiene.

<u>LEMA 1.</u>

Si en el modelo directo se tiene que A_{11}, A_{22} son invertibles con inversa causal entonces se puede derivar un modelo realimentado.

<u>DEMOSTRACION.</u>

Para el modelo directo se tendrá (Fig. 1)

$$y = A_{11}\, W_1 + A_{12}\, W_2$$
$$u = A_{21}\, W_1 + A_{22}\, W_2\ .$$

Si A_{11}, A_{22} son invertibles, se cumple

$$W_1 = A_{11}^{-1} (y - A_{12} W_2)$$

$$u = A_{21} A_{11}^{-1} y + (A_{22} - A_{21} A_{11}^{-1} A_{12}) W_2$$

Es decir

$$H = A_{21} A_{11}^{-1}$$

$$K = A_{22} - A_{21} A_{11}^{-1} A_{12}$$

Como A_{11}^{-1} es causal por hipótesis y todos los demás también lo son, se deduce que H y K son operadores causales.

En forma similar se obtendrá

$$F = A_{12} A_{22}^{-1}$$

$$G = A_{11} - A_{12} A_{22}^{-1} A_{21} .$$

Esto completala demostración. Nótese que $H = 0$ (lazo abierto) si y solo si $A_{21} = 0$.

El interés de este resultado es que existe un caso especial de factorización espectral para el cual las condiciones del Lema son aplicables.

DEFINICION 1:

Un operador $Q_\zeta > 0$ admite una factorización radical si existe un operador L estrictamente causal y un operador M>0 sin memoria (estático) tal que

$$Q_\zeta = (I + L) M (I + L^*)$$

La factorización radical tiene propiedades que la hacen especialmente adecuada para la modelación. Constituye la extensión del concepto de densidad espectral normalizada de [6]. Además de existir eficientes algoritmos para determinarla [8] , se dan sin demostración las siguientes propiedades [7] que justifican su utilidad.

P1 : La factorización radical es única.

P2 : Dado que L es estrictamente causal, I+L es invertible y con inversa causal.

Al aplicaresta definición al proceso $\zeta = (y, u)$ en H^2, se puede escribir

$$I + L = \begin{bmatrix} (I + L_{11}) & L_{12} \\ L_{21} & (I + L_{22}) \end{bmatrix} \quad M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix}$$

Llamando $A_{11} = I + L_{11}$, $A_{22} = I + L_{22}$ se ve que ambos son causales con inversa causal. Siempre es posible plantear un modelo realimentado en este caso. El operador M representará la covariancia del proceso $\eta = (n_1, n_2)$ de señales externas.

4.- REALIMENTACION Y PSEUDO REALIMENTACION.

En la sección anterior se han planteado dos posibles modelos para representar la interacción entre procesos estocásticos. El modelo directo supone que ambos procesos tienen "causas" comunes mientras que el modelo realimentado supone una relación causa/efecto,en general bidireccional, entre los procesos. Ambos pueden derivarse de la misma información.

En este punto se hace necesario recalcar la diferecia conceptual entre un modelo matemático y el objeto representado por dicho modelo. Específicamente se tiene un ente (físico) en el cual se han medido dos variables, modeladas a su vez como los procesos estocásticos (y,u). Es a este ente físico al cual se le quiere asociar un modelo basado sólo en las estadísticas de segundo orden. Para tratar de mantener la diferencia se empleará el término sistema para referirse a este ente físico.

Debe ser claro que a partir de la información disponible no se puede hacer ninguna afirmación sobre el sistema. Es más, bajo condiciones simples, un sistema con estructura realimentada puede ser representado por un modelo directo. Es también posible que un sistema en el cual no hay realimentación, por efecto de la modelación, aparezca como poseyendo una estructura realimentada.

Para poder avanzar es necesario hacer una suposición adicional. En este caso se supondrá a priori que existe una relación causa/efecto entre las variables (y,u).

El sistema en el cual se han medidc los procesos estocásticos se supone con una estructura realimentada como la mostrada en la Figura 2. El problema que se debe resolver en este caso no es solo el de determinar un modelo de la misma forma sino que determinar también bajo que condiciones se puede establecer que en el sistema no hay realimentación ($H = 0$) de modo que la relación causa efecto es completamente unidireccional.

En este contexto, el concepto de ausencia de realimentación jugará un papel destacado. Antes de plantearlo formalmente es necesario recalcar que está completamente basado en las estadísticas de segundo y es por lo tanto verificable con la información dada e independiente de cualquier estructura que pueda tener el sistema. Además, a diferencia de las definiciones que se han publicado ([4], [5], [6]) la que aquí se dá es válida para procesos estocásticos no estacionarios.

DEFINICION 2:

No hay realimentación en sentido fuerte del proceso y sobre el proceso u si y solo si la covariancia Q_ζ del proceso $\zeta = (y, u)$ admite una factorización espectral $A A^*$ en H^2 para la cual se cumpla

i) $A_{21} = 0$

ii) A_{12} estrictamente causal.

Por condición de factorización espectral la inversa de A debe ser también causal. Como en el caso de ausencia de realimentación se tiene $A_{21}=0$ se puede ver facilmente que en este caso A_{11} y A_{22} son ambos con inversa causal. Utilizando el Lema 1 se concluye que existe un modelo realimentado para las interacciones entre el par (y, u).

Siguiendo el desarrollo del Lema 1 se ve que en el caso de ausencia de realimentación en sentido fuerte se tendrá un modelo realimentado para el cual F es estrictamente causal y $H = 0$. Es decir el modelo es de lazo abierto y claramente unidireccional. Las señales de ruido externas (n_1, n_2) (ver Figura 2) son independientes entre si. De acuerdo a este modelo se puede concluir que la variable u es "causa" con respecto a la variable y. Mas específicamente que la variable y se puede variar mediante manipulación de la variable u pero no a la inversa.

El concepto sería claramente inutil si con la misma información, la covariancia Q_ζ, se puede obtener también la conclusión opuesta. Esta posibilidad queda descartada con el siguiente resultado.

LEMA 2.

Dado el par de procesos $\zeta = (y,u)$ con covariancia $Q_\zeta > 0$, las siguientes proposiciones son equivalentes

i) Los procesos son estadísticamente independientes

ii) No hay realimentacion (en sentido fuerte) del proceso y sobre el proceso u, ni del proceso u sobre y.

DEMOSTRACION.

i) $\Rightarrow$ ii) : Esto es muy fácil de ver ya que en este caso $Q_{yu} = Q_{uy} = 0$. El operador Q_ζ es por lo tanto diagonal. Cada uno de los elementos Q_y, Q_u es positivo definido. Aplicando factorización espectral a cada uno de ellos se obtiene una factorización $A A^*$ para Q_ζ en la cual $A_{12} = A_{21} = 0$. El resultado sigue inmediatamente aplicando la definición.

ii) $\Rightarrow$ i) : En este caso se tendrán en general dos factorizaciones $A A^*$, $B B^*$ para Q_ζ. En una de ellas se tendrá $A_{21} = 0$, A_{12} estrictamente causal (ausencia de realimentación de y sobre u). En la otra factorización se tendrá en cambio $B_{12} = 0$, B_{21} estrictamente causal, indicando ausencia de realimentación de u sobre el proceso y.

Dado que $A A^* = B B^* = Q_\zeta$ debe cumplirse $B = A N$, $N N^* = I$, N sin memoria (sección 3). Usando representación matricial.

$$\begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$$

De aquí se obtiene

$$0 = A_{11} N_{12} + A_{12} N_{22}$$

$$B_{21} = A_{22} N_{21}$$

$$N_{21} = B_{21} A_{22}^{-1}$$

$$N_{12} = - A_{11}^{-1} A_{12} N_{22}$$

Como por hipótesis B_{21} y A_{12} son estrictamente causales y todos los otros operadores son causales, debe cumplirse que tanto N_{21} como N_{12} son estrictamente causales y por lo tanto su componente sin memoria debe ser nula. Como al mismo tiempo N_{12}, N_{21} son sin memoria debe cumplirse

$$N_{12} = N_{21} = 0$$

De aquí usando algebra elemental se deduce $A_{12} = 0$, $B_{21} = 0$. Por lo tanto Q_ζ es diagonal y los procesos son estadísticamente independientes. El lema queda completamente demostrado.

Al mismo tiempo el lema deja en evidencia (eligiendo N adecuadamente) que aún cuando se tenga ausencia de realimentación en sentido fuerte del proceso y sobre el proceso u, se puede obtener una factorización espectral $B B^*$ en la cual tanto B_{12} como B_{21} son no nulos. Por lo tanto se pueden también obtener modelos realimentados con $F \neq 0$ y $H \neq 0$ (Formalmente el único caso excluido es $F = 0$, $H \neq 0$). Debe notarse sin embargo que en el resultado anterior no se ha empleado la condición $N N^* = I$. Esta condición resulta vital para el siguiente resultado.

LEMA 3.

Dado el par de procesos $\zeta = (y,u)$ con $Q_\zeta > 0$ si no hay realimentación en sentido fuerte de y sobre u entonces no existe un modelo realimentado que satisfaga las condiciones (ver Figura 2).

i) Los operadores F,G,H,K son causales

ii) F ó H (o ambos) es estrictamente causal

iii) Las entradas externas $\eta = (n_1, n_2)$ son secuencias blancas mutuamente independientes

iv) $H \neq 0$.

DEMOSTRACION.

Cualquier modelo realimentado satisface (Figura 2)

$$y = Fu + Gn_1$$

$$u = Hy + Kn_2 \ .$$

Dado que F (o H) es estrictamente causal entonces FH y HF son también operadores estrictamente causales. Por lo tanto el operador I-FH (I-HF) es invertible y con inversa causal. Se puede entonces resolver el sistema de ecuaciones obteniendose

$$y = (I-FH)^{-1} (FK\, n_2 + G\, n_1)$$

$$u = (I-HF)^{-1} (K\, n_2 + HG\, n_1).$$

Sin perder generalidad se puede considerar que $\eta = (n_1, n_2)$ satisface $Q_\eta = I$ en H^2. De aquí sigue que definiendo

$$B_{11} = (I - FH)^{-1} G$$

$$B_{12} = (I - FH)^{-1} FK$$

$$B_{21} = (I - HF)^{-1} HG$$

$$B_{22} = (I - HF)\ K;$$

se obtiene una factorización espectral $B B^*$ para el proceso conjunto $\zeta = (y,u)$ en la cual ya sea B_{12} o B_{21} (o ambos) son operadores estrictamente causales.

Por hipótesis hay ausencia de realimentación del proceso y sobre el proceso u. De modo que Q_ζ tambien admite una factorización espectral $A A^*$ con $A_{21} = 0$, A_{12} estrictamente causal. Se debe cumplir (para algún N sin memoria)

$$B = A N$$

$$N N^* = I.$$

Desarrollando las expresiones como en el lema anterior se tiene

$$B_{12} = A_{11} N_{12} + A_{12} N_{22}$$

$$B_{21} = A_{22} N_{21} \ .$$

Se ve que la condición B_{21} estrictamente causal implica $N_{21} = 0$, al igual que en lema anterior. Por otra parte dado que

$$N_{12} = A_{11}^{-1} (B_{12} - A_{12} N_{22}),$$

la condición B_{12} estrictamente causal implica que N_{12} debe ser cero.

Por otra parte, desarrollando la condición $N N^* = I$ entérminos de los elementos N_{ij} se obtiene fácilmente que $N_{12} = 0$ si y solo si $N_{21} = 0$. De modo que B_{12} o B_{21} estrictamente causal implican N diagonal.

Ahora bien si N es diagonal, de la expresión $B = AN$ se ve que A triangular superior ($A_{21} = 0$) implica B triangular superior ($B_{21} = 0$). Por lo tanto $H = 0$. (Nota: En estricto rigor se obtiene $HG = 0$ pero se puede establecer que G debe ser invertible para que $Q_\zeta > 0$).

Debe notarse que las tres primeras condiciones del lema son normalmente suposiciones que se plantean en el proceso de modelación. Para dejar este punto más claro se tiene:

COROLARIO:

Si se sabe que el sistema que genera los procesos $\zeta = (y,u)$ satisface condiciones i), ii), iii) del Lema 3 y se determina ausencia de realimentación en sentido fuerte de y sobre u entonces necesariamente deba tenerse $H = 0$.

Como consecuencia de este resultado se tiene una condición para determinar si un sistema está en lazo abierto ($H = 0$) o no.

Desgraciadamente existen sistemas a los cuales se desearía clasificar como de lazo abierto y que sin embargo no satisfacen el criterio de ausencia de realimentación en sentido fuerte.

Considérese el sistema mostrado en la Figura 3. El operador F es causal y los procesos estocásticos α y $\beta = u$ son secuencias blancas independientes entre si y de covariancia unitaria.

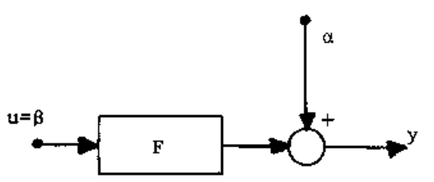

Figura 3: Sistema en Lazo Abierto

Se puede demostrar sin mucho esfuerzo que para $\zeta = (y,u)$ se cumple

$$Q_\zeta = \begin{bmatrix} I & F \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ F^* & I \end{bmatrix}$$

Como F es causal se tiene de hecho una factorización espectral inmediata para Q_ζ. Como F puede tener parte sin memoria (estática), para esta factorización no se cumplen las condiciones para ausencia de realimentación en sentido fuerte. Si se supone que existe una factorización para la cual se cumplen las condiciones, debería tenerse

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \begin{bmatrix} I & F \\ 0 & I \end{bmatrix} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix},$$

Con A_{12} estrictamente causal. Sin embargo es claro que $A_{21} = 0$ implica $N_{21} = 0$ y por consiguiente $N_{12} = 0$ de modo que

$$A_{12} = F N_{22} .$$

Como N_{22} es estático, la condición A_{12} estrictamente causal se cumple si y solo si F lo es. Así se tiene que el sistema de la figura 3 puede no cumplir los requisitos para ausencia de realimentación en sentido fuerte.

Esta anomalia hace necesario introducir un concepto de ausencia de realimentación en sentido débil. Específicamente se tiene.

DEFINICION 3.

Dados los procesos $\zeta = (y,u)$ con $Q_\zeta > 0$ se dice que hay ausencia de realimentación en sentido débil del proceso y sobre el proceso u si y solo si existe una factorización espectral $A A^*$ para Q_ζ y tal que $A_{21} = 0$.

La única diferencia entre esta definición y la anterior es la eliminación del requisito que A_{12} sea estrictamente causal. Esto permite considerar al sistema de la figura 3 como sin realimentación en sentido débil. Claramente este es un concepto más amplio ya que todo par sin realimentación en sentido fuerte lo es también en sentido débil pero no a la inversa.

Debe notarse que ambas definiciones son muy complicadas de aplicar en la forma dada ya que piden la existencia de alguna factorización espectral. En este sentido no son constructivas.

El siguiente resultado resuelve este problema para ausencia de realimentación en sentido fuerte.

TEOREMA 1.

Dados los procesos $\zeta = (y,u)$ con $Q_\zeta > 0$ las siguientes proposiciones son equivalentes

i) Hay ausencia de realimentación en sentido fuerte de y sobre u.

ii) En la factorización radical de Q_ζ se tiene (sección 3) L triangular superior ($L_{21} = 0$) y M diagonal ($M_{12} = M_{21} = 0$).

DEMOSTRACION.

ii) ⇒ i) Con las condiciones dadas se tiene un operado L estrictamente causal y M sin memoria tales que

$$Q_\zeta = (I + L)\ (M)\ (I + L^*)$$

$$= \begin{bmatrix} (I + L_{11}) & L_{12} \\ 0 & (I + L_{22}) \end{bmatrix} \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix} \begin{bmatrix} (I + L_{11}^*) & 0 \\ L_{12}^* & (I + L_{22}^*) \end{bmatrix}.$$

Como $M_{11}>0$, $M_{22}>0$ se puede escribir

$$\begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix} = \begin{bmatrix} N_{11} & 0 \\ 0 & N_{22} \end{bmatrix} \begin{bmatrix} N_{11}^{*} & 0 \\ 0 & N_{22}^{*} \end{bmatrix}$$

Con N_{ii} operador estático.

Por lo tanto definiendo

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \left[\begin{array}{c|c} (I + L_{11})\, N_{11} & L_{12}\, N_{22} \\ \hline 0 & (I + L_{22})\, N_{22} \end{array}\right]$$

Se ve que A A* es una factorización espectral que satisface la definición de ausencia de realimentación en sentido fuerte.

i) $\Rightarrow$ ii) Como A es triangular, A_{11} (A_{22}) es invertible con inversa causal. Separando A_{11}^{-1}, A_{11} en sus componentes sin memoria y estrictamente causal se puede probar que la parte sin memoria $[A_{11}]_z$, de A_{11} debe ser no singular para satisfacer $A_{11}\, A_{11}^{-1} = I$. Por lo que es posible definir

$$L_{11} = [A_{11}]_{\underline{c}}\ [A_{11}]_z^{-1}$$

$$N_{11} = [A_{11}]_z$$

Claramente L_{11} es estrictamente causal y satisface

$$A_{11} = (I + L_{11})\, N_{11}$$

En forma similar se establece la existencia de L_{22}, N_{22} y finalmente para L_{12} se tiene

$$L_{12} = A_{12}\, N_{22}^{-1}$$

y también es estrictamente causal ya que A_{12} lo es. La existencia de la factorización radical está establecida. El teorema queda completamente demostrado.

Una de las razones principales que hacen importante este resultado es que como se ha mencionado anteriormente, la factorización radical es única y se conocen algoritmos eficientes para calcularla.

Este teorema dá entonces un procedimiento constructivo para verificar la ausencia de realimentación en sentido fuerte.

En forma muy parecida es posible caracterizar la ausencia de realimentación en sentido débil. Por ser la demostración bastante parecida se omite.

TEOREMA 2.

Dados los procesos $\zeta = (y,u)$ con $Q_\zeta>0$ las siguientes proposiciones son equivalentes:

i) Hay ausencia de realimentación en sentido débil de y sobre u.

ii) En la factorización radical de Q_ζ se tiene L triangular superior ($L_{21} = 0$).

El hecho que el operador M pueda tener componentes fuera de la diagonal se puede interpretar también como una correlación entre las señales de ruido (n_1, n_2).

Sin información adicional no se tiene elementos para preferir un modelo con H = 0, F causal y señales externas independientes sobre un modelo con H = 0, F estrictamente causal y señales externas correlacionadas. Esto permite clarificar algunas confusiones existentes en la literatura sobre "sistemas sin realimentación".

La caracterización dada por el Teorema 2 permite obtener algunos resultados adicionales realmente interesantes.

Dado que todo modelo causal con H = 0 y entradas externas independientes implica y es implicado por $L_{21} = 0$ se puede plantear :

COROLARIO 1:

Si en la factorización radical de Q_ζ se tiene $L_{21} \neq 0$ entonces cualquier modelo causal realimentado y con entradas no correlacionadas requiere un operador $H \neq 0$. Es decir requiere realimentación del proceso y sobre el proceso u.

Se pueden plantear aún otros resultados referentes a la estructura del sistema que genera a los procesos $\zeta = (y,u)$ indicando que la factorización radical contiene una cantidad enorme de información estructural y es el camino lógico para analizar interacciones. Por razones de espacio no es posible considerar más implicaciones y se prefiere abordar otra caracterización adicional de los conceptos de ausencia de realimentación.

5.- RELACION CON EL PROBLEMA DE ESTIMACION ESTOCASTICA.

Se mencionó en la introducción que el problema de modelar interacciones se ha planteado también en términos del problema de estimación.

En esta sección se anliza brevemente esta conexión y se extienden resultados que existen para el caso estacionario ([1] , [7]).

Dados dos procesos estocásticos se puede plantear el problema de estimar uno de ellos (y) basado en algún conocimiento del otro (u). Usando estadísticas de segundo orden se tiene el problema de determinar un sistema lineal (operador) D, tal que

$$E\ \{\|\, y - Du \,\|^2\} = \min.$$

Cuando el operador D debe satisfacer una condición de "Realizabilidad física" se tiene el problema de Wiener. En términos de espacios de resolución ésta es una condición de causalidad.

Designando por $e = y-Du$ se puede definir

$$J(D) = E\,\{\|e\|^2\} = \mathrm{tr}\,\{Q_e\}$$

$$= \mathrm{tr}\,\{Q_y - Q_{yu}D^* - DQ_{uy} + D\,Q_u D^*\}$$

$$J\,(D + \delta D) = J(D) + \mathrm{tr}\,\{\delta\, D\, Q_u\ \delta\, D^*\} +$$

$$\mathrm{tr}\,\{\delta\, D\, Q_u\, D^* + D\, Q_u \delta\, D^* - Q_{yu}\delta D^* - \delta DQ_{uy}\}.$$

Para $\delta J = J(D + \delta D) - J\,(D)$ se cumple

$$\delta J = \mathrm{tr}\,\{\delta D\ Q_u\ \delta D^*\} + \mathrm{tr}\,\{\delta\, D\,[Q_u D^* - Q_{uy}]\,\}$$

$$+\ \mathrm{tr}\,\{\,[D\, Q_u - Q_{yu}]\ \ \delta\, D^*\,\}\,.$$

El término $\mathrm{tr}\{\delta D\, Q_y\, \delta D^*\}$ es siempre no negativo. La condición necesaria y suficiente para minimizar J es por lo tanto

$$\mathrm{tr}\left\{\left[\hat{D}\, Q_u - Q_{yu}\right]\delta D^*\right\} = 0.$$

Esta relación debe cumplirse para todo δD en la colección de aquellos considerados admisibles.

Si no hay ninguna restricción en la elección de D entonces δD debe ser cualquier operador. La solución que se obtiene en este caso es

$$\hat{D}_o = Q_{yu}\, Q_u^{-1}.$$

Esta solución $\hat{D}_o$ será en general no causal.

Si se exige que D deba ser causal entonces δD debe ser restringido a todos los operadores causales solamente. La solución en este caso está dada por la solución de las ecuaciones de causalidad

$$\left[\hat{D}_c Q_u\right]_c = \left[Q_{yu}\right]_c$$

$$\hat{D}_c = \left[\hat{D}_c\right]_c.$$

De la misma manera si se exige que D debe ser estrictamente causal se obtiene una solución $\hat{D}_{\underline{c}}$.

En general se cumplirá que las tres soluciones, $\hat{D}_o$, $\hat{D}_c$, $\hat{D}_{\underline{c}}$ son diferentes. No es sorprendente que ellospuedan coincidir en algunos casos. Lo que es interesante es que estos casos están directamente ligados a la ausencia de realimentaciones como lo establece el siguiente resultado

LEMA 3.

La siguientes condiciones son equivalentes

a) $\hat{D}_o = \hat{D}_{\underline{c}}$ (/ $\hat{D}_o = \hat{D}_c$)

b) No hay realimentación en sentido fuerte (/débil) de sobre u.

DEMOSTRACION:

b) $\Longrightarrow$ a) Si no hay realimentación de y sobre u se tiene un modelo-F en lazo abierto (H=0).

$$y = Fu + Gn_1$$

$$u = Kn_2 .$$

Los procesos n_1, n_2 son ruidos blancos independientes.

$$Q_{yu} = FQ_u .$$

De modo que

$$\hat{D}_o = F.$$

Si la ausencia de realimentación es en sentido fuerte entonces F es estrictamente causal. Si se habla en sentido débil F es sólo causal. La proposición sigue inmediatamente

a) $\Longrightarrow$ b) Dado que $D_o = Q_{yu}Q_u^{-1}$ y es (estrictamente /) causal se puede escribir

$$y = D_o u + \alpha$$

$$u = \beta$$

con $Q_{u\alpha} = 0$. Los operadores G,K se determinan por el procedimiento de blanquear los procesos α y β.

El lemma queda así establecido. Por primera vez se demuestra esta equivalencia para el caso de ausencia de realimentación en sentido débil.

Otras caracterizaciones son también posibles. Sin embargo se consideran de menor importancia y se omiten por razones de espacio.

6.- CONCLUSIONES.

Se ha estudiado el problema de analizar interacciones entre procesos estocásticos basándose en estadísticas de segundo orden.

El enfoque de espacios de resolución y teoría de causalidad permite extender sin ninguna dificultad resultados existentes sólo para procesos estacionarios.

Se obtienen resultados valiosos en relacionar ausencia de realimentación débil y fuerte. Queda además completamente clara la limitación del procedimiento demostrando que existen situaciones que no se pueden discriminar.

De particular importancia es el resultado que sistemas en lazo abierto podrían ser considerados como realimentados en el sentido fuerte. Esto haría aparecer una pseudo realimentación que puede distorsionar completamente el proceso de modelación ya que genera una relación causa/efecto ficticia.

Finalmente es interesante que la teoría de causalidad tiene aplicación y produce resultados muy poderosos en la determinación de relación causa/efecto entre procesos estocásticos.

REFERENCIAS.

[1] Sims, C.A., "Money, Income and Causality", Amer. Econ. Rev., Vol. 62, pp 540-552, 1972.

[2] Gersch, W., "Causality or Driving in Electrophysiological Signal Analysis", Hath. Biosci., Vol 14, pp 177-196, 1972.

[3] Good hart D., et al., "Money, Income and Causality: The UK Experience", Amer. Econ. Rev., 1976.

[4] Caines, P.E., Chan, C.W., "Feedback Between Stationary Stochastic Processes". IEEE Trans. Aut. Cont., Vol AC-20, N°4, pp 498-508, Aug. 1975.

[5] Caines P.E., "Weak and Strong Feedback Free Processes", IEEE Trans. Aut. Cont. pp 737-739, Oct. 1976.

[6] Gevers, M.R., Anderson, B.D.O. "On Jointly Stationary Feedback Free Stochastic Processes", IEEE Trans. Aut. Cont., Vol AC-27, N°2, pp 431-436, April 1982.

[7] Feintuch, A., Saeks, R., "System Theory A Hilbert Space Approach", Academic Press, 1982.

[8] Porter, W.A., Aravena, J.L., "Algorithms for Fast Multidimensional Signal Extraction", IEEE Trans. Circ. Syst. (to appear 1983).

[9] De Santis, R.M., Porter, W.A., "Operator Factorization in Partially Ordered Hilbert Resolution Spaces" Math. Syst. Theo (to appear 1983).

[10] Saeks, R. "Resolution Space Operatos and Systems", Lectures Notes in Economics and Mathematical Systems, Springer Verlag, New York, 1973.

ARAVENA LUMAN, JORGE. Ingeniero Civil Electricista, U. de Chile, Ph.D., Computadores, Información y Control, U. de Michigan.
El Dr. Aravena comenzó su actividad académica en 1967 como Miembro del Departamento de Electricidad, Universidad de Chile, donde llegó a ser Jefe del Grupo de Control Automático y Coordinador de Investigación.
Como Miembro y Jefe de PROSCOM, Universidad de Concepción, participó en proyectos de automatización con computador para la gran minería del cobre.
Como integrante del Departamento de Ingeniería Eléctrica y Computadores, Universidad de Louisiana, realizó durante cuatro años docencia e investigación en control y ciencia de sistemas.
Desde Julio 1982, el Dr. Aravena es Investigador y Docente en el Departamento de Ingeniería Eléctrica, Universidad de Santiago.
El Dr. Aravena se encuentra activamente envuelto en investigación relacionada con conceptos básicos de sistemas dinámicos y es autor de numerosos artículos en revistas técnicas internacionales y congresos especializados.