

REDUCCIÓN DE LA DIMENSIÓN DE SERIES TEMPORALES HEMODINÁMICAS MEDIANTE MODELOS GRÁFICOS

R. Fried

Departamento de Estadística. Universidad Carlos III, Madrid

Antecedentes y objetivos

En la medicina intensiva la monitorización de cada paciente incluye la observación casi continua de una gran cantidad de variables fisiológicas: diferentes presiones sanguíneas, pulso, temperatura, etc. Estas series temporales hemodinámicas proporcionan información crucial sobre el estado del paciente crítico. La detección de cambios en sus trayectorias es de especial interés práctico, ya que nos informan de posibles complicaciones en el estado del paciente y nos permiten valorar la eficacia de la terapia aplicada. La elevada dimensionalidad de las series observadas y la necesidad de tomar decisiones rápidas, ponen en evidencia las limitaciones de los seres humanos, cuyas decisiones equivocadas pueden producir consecuencias fatales.

La extracción e interpretación correcta de la información de todo este torrente de datos resulta muy difícil de lograr con sólo la experiencia profesional del médico. En consecuencia, necesitamos construir sistemas fiables que apoyen las decisiones del médico intensivista y disponer de métodos que permitan reducir la dimensión de las variables. Los médicos, habitualmente, utilizan sólo un pequeño número de variables de entre todas las disponibles. Como estos procedimientos de selección son subjetivos e ineficaces debido a la gran cantidad de información relevante desperdiciada, es importante conocer la información perdida y analizar las relaciones de todas las variables entre sí.

Métodos

Como alternativa a un procedimiento de selección de variables podemos utilizar componentes principales o construir un modelo factorial como el propuesto en Peña y Box (2002) que explique una mayor cantidad de variabilidad con el mismo número de variables (Gather et al., 2001). En este caso tenemos algunos problemas en la interpretación de las variables latentes y también resulta difícil aplicar rotaciones en presencia de interacciones dinámicas. Para facilitar la interpretación de las variables latentes podemos incluir información sobre las relaciones conjuntas de las variables entre sí (Fried y Didelez, 2002).

Una posible fuente de información puede ser la construcción de un modelo gráfico como el propuesto por Brillinger (1996) y Dahlhaus (2000). Este tipo de modelos

permite una visualización simple y comprensible de las relaciones entre las distintas variables: las variables se representan por vértices y las correlaciones parciales por aristas. Para su construcción, estimamos las coherencias espectrales parciales, que miden las correlaciones parciales en el dominio de frecuencias. Podemos interpretar la ausencia de aristas entre las variables utilizando propiedades markovianas: la propiedad markoviana dos a dos, que utilizamos para la construcción del modelo, nos dice que la ausencia de una arista entre dos variables implica que dichas variables no tienen relación lineal dados los efectos lineales del resto de variables. Dahlhaus (2000) demuestra que, si la matriz de las densidades espectrales es regular para todas las frecuencias, la anterior propiedad implica la markoviana global, que es mucho más cómoda de utilizar: si en el modelo construido dos grupos de variables A y B están separados por otro grupo S (que significa que siempre y cuando queramos pasar de A a B debemos visitar S), entonces no hay relación lineal entre las variables de A y de B dados los efectos de las pertenecientes a S. Por lo tanto estos modelos generalizan los modelos gráficos, bien conocidos para observaciones independientes, a la situación dinámica en que las observaciones forman series temporales multivariantes.

En este escenario, la selección específica de un modelo gráfico para una serie es un problema básico. Podemos estimar las coherencias espectrales parciales y contrastar si cada una de ellas es simultáneamente cero o no. En esta etapa, al igual que en la selección de covariables en regresión múltiple, los contrastes pueden ser poco informativos debido a la multitud de interacciones que necesitamos estimar. Por tanto, una estrategia más avanzada para seleccionar el modelo resulta de especial utilidad.

Resultados

Primero verificamos algunas propiedades de los modelos gráficos para el caso dinámico que son útiles para seleccionar un modelo en etapas Fried y Didelez (2003). Bajo la hipótesis de que conocemos el modelo correcto, las aristas ausentes entre un grupo de variables son válidas también en un modelo gráfico reducido con sólo estas variables. La excepción es que necesitamos incluir aristas entre parejas de variables si ambas tienen correlación parcial no nula con una o varias de las variables excluidas. Aún más fuerte, la coherencia espectral parcial entre dos variables en el modelo reducido es la misma que en el modelo general con todas las variables, siempre que ambas variables no estén conectadas con ninguna de las variables excluidas. A partir de estas propiedades podemos seleccionar un modelo gráfico por etapas. Primero construimos un modelo en base a los contrastes simultáneos para todas las correlaciones parciales. Luego verificamos la ausencia y la presencia de aristas utilizando separaciones en el modelo concreto seleccionado.

Aplicamos esta estrategia para construir un modelo gráfico para el sistema hemodinámico de cada uno de veinte pacientes estudiados. En todos los modelos seleccionados

encontramos al final los mismos grupos de variables con correlaciones parciales muy altas entre sí. Además, estos grupos tienen una interpretación intuitiva que es razonable para los médicos porque se presentan a través de presiones arteriales, arterial pulmonares, pulso y frecuencia cardíaca. Estos resultados sugieren una representación de cada grupo a través de una variable latente. Y finalmente, señalamos que con este método conseguimos explicar mucha mayor variabilidad que con un procedimiento de selección de variables.

Conclusiones

Los modelos gráficos sugeridos por Brillinger (1996) y Dahlhaus (2000) son métodos muy útiles para analizar relaciones lineales entre series temporales multivariantes que además tienen propiedades parecidas al caso estático: en la aplicación a series temporales hemodinámicas, los resultados coinciden con las relaciones usualmente conocidas. Por lo tanto, esperamos obtener información relevante también en aplicaciones a series temporales multivariadas en las que no se dispone de mucho conocimiento y experiencia.

Agradecimientos

Quiero expresar mi más profunda gratitud a Juan Ferrándiz, llamado “Gudo”, quien me familiarizó con los modelos gráficos mientras disfrutaba de su hospitalidad y amistad en Valencia durante el verano de 1997.

Referencias

- Brillinger, D. (1996). Remarks concerning graphical models for time series and point processes. *Revista de Econometría* 16:1–23.
- Dahlhaus, R. (2000). Graphical interaction models for multivariate time series. *Metrika* 51:157–172.
- Fried, R. y Didelez, V. (2002). Latent variable analysis and partial correlation graphs for multivariate time series. T. Rep. 6/2003, SFB 475, Universidad de Dortmund, Alemania.
- Fried, R. y Didelez, V. (2003). Decomposability and selection of graphical models for multivariate time series. *Biometrika* 90:251–267.
- Gather, U., Fried, R., Lanius, V. y Imhoff, M. (2001). Online monitoring of high dimensional physiological time series - a case study. *Estadística* 53:259–298.
- Peña, D. y Box, G. E. P. (2002). Identifying a simplifying structure in time series. *Journal of the American Statistical Association* 82:836–843.