

*Pronósticos e Intervención
en Sistemas Dinámicos de Flujo*

Liliana Figueroa Quiroz

Seminario Aleatorio

I T A M

Marzo 2005

Motivación

Redes Dinámicas de Flujo

Modelos Estadísticos de Flujos en Redes

Redes de Flujo vs. Redes Bayesianas

Modelo Lineal Dinámico para Flujos en Redes

Intervención en Modelos de Flujos en Redes

Comentarios Finales

Motivación

A través de una red estructurada, fluye material en el tiempo.

petróleo crudo

Contamos con distintas fuentes de información.

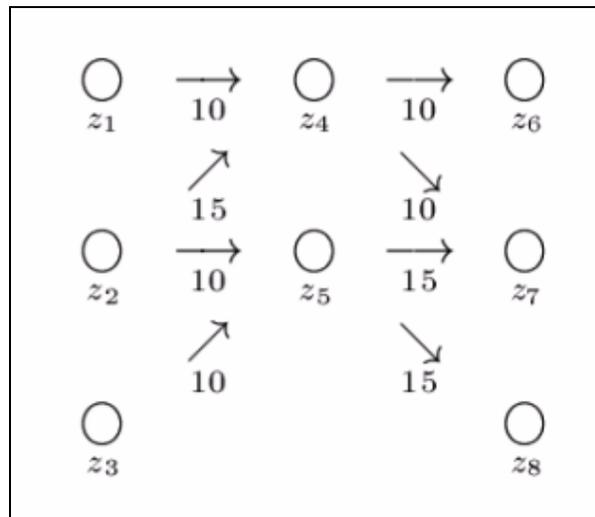
OPEP, agencias

La red está sujeta a eventos que modifican los flujos.

Guerra del Golfo

Queremos construir un modelo de pronóstico que sea adaptable bajo ciertas manipulaciones.

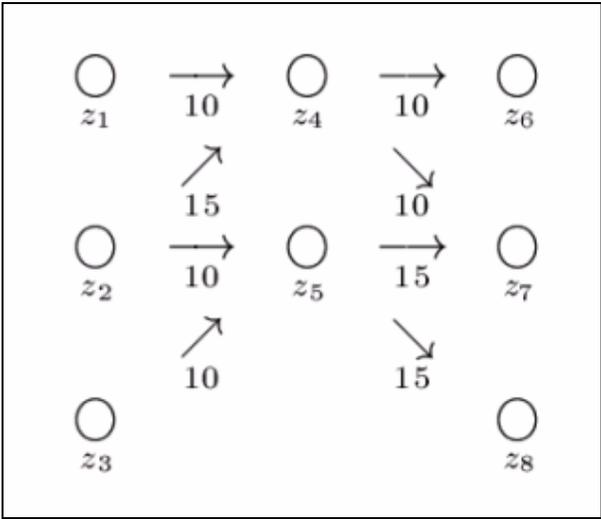
Redes Dinámicas Jerárquicas de Flujos



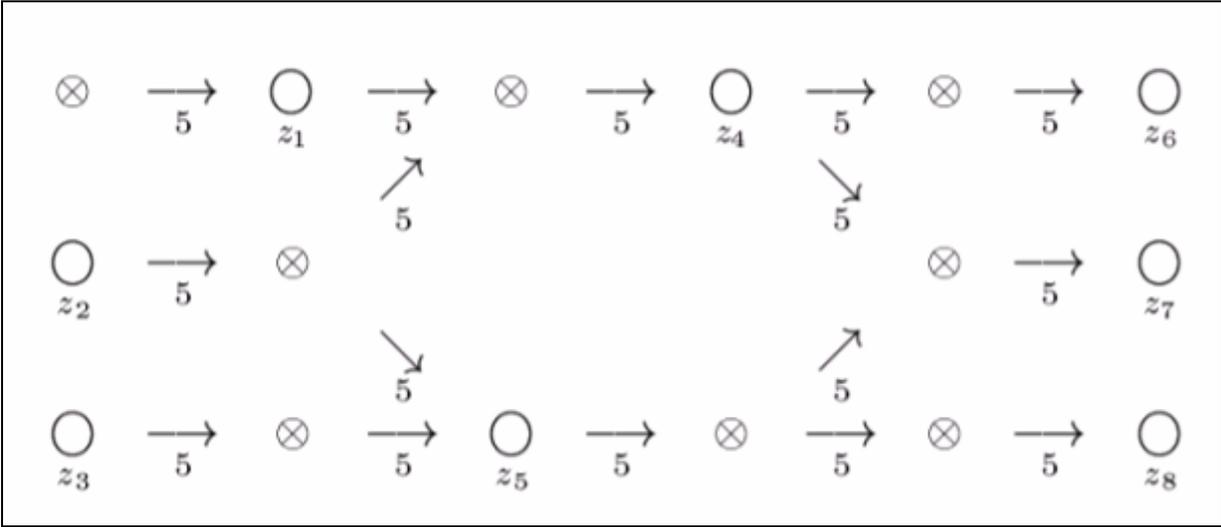
- conjunto de nodos con una partición en m clases C^l
- es posible transferir material sólo de C^l a C^{l+1} , $l=1, \dots, m-1$
- los nodos en C^1 no reciben material de otros nodos
- los nodos en C^m no transfieren material a otros nodos
- se cumple la propiedad de la conservación de la materia
- el tiempo de transferencia de material entre nodos no es necesariamente igual

Bajo ciertas condiciones, se puede crear una red de flujos que “emule” a la red original en la que los tiempos de transferencia sean iguales: **Red Recursiva de Flujos (RFN)**

Redes Recursivas de Flujos



Red Jerárquica de Flujos en Equilibrio



Red Recursiva de Flujos

Modelos Estadísticos de Flujos en Redes

Motivados por problemas de tráfico de vehículos:

Whittaker, Garside & Lindveld (1997)

Tebaldi & West (1998)

Whitlock & Queen (2000)

Zohar & Geiger (2002)

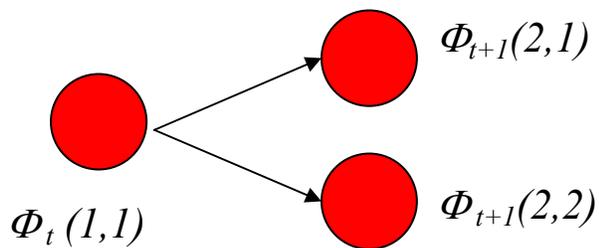
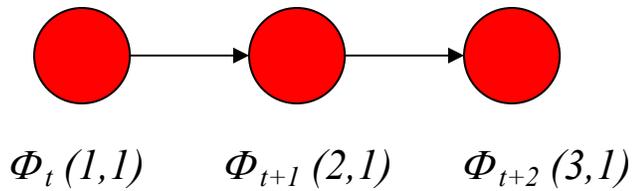
No son completamente efectivos al no

- ✚ modelar los datos en el tiempo y el espacio
- ✚ distinguir observaciones de los procesos reales
- ✚ contar con una descomposición útil de los procesos
- ✚ definir una metodología de intervención

Redes de Flujo vs. Redes Bayesianas

Las redes de flujo son acíclicas.

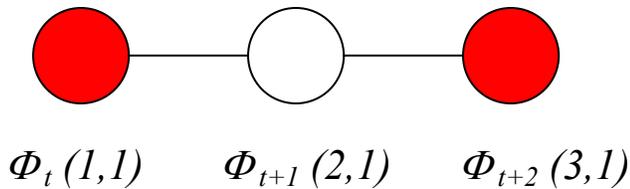
¿Podemos asumir que son Redes Bayesianas (Pearl, 1988)?



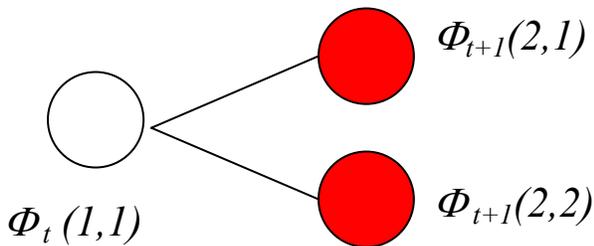
Redes de Flujo vs. Redes Bayesianas

Las redes de flujo son acíclicas.

¿Podemos asumir que son Redes Bayesianas (Pearl, 1988)?



$$\Phi_t(1,1) \perp\!\!\!\perp \Phi_{t+2}(3,1) \mid \Phi_{t+1}(2,1)$$



$$\Phi_{t+1}(2,1) \not\perp\!\!\!\perp \Phi_{t+1}(2,2) \mid \Phi_t(1,1)$$

La propiedad de la conservación de la materia exige que

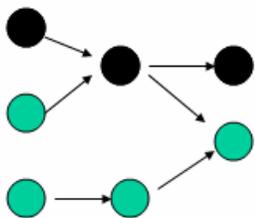
$$\Phi_{t+1}(2,1) + \Phi_{t+1}(2,2) = \Phi_t(1,1)$$

Las propiedades de transmisión de información no pueden asumirse en general para la transferencia de material.

Modelo Lineal Dinámico para Flujos en Redes

El modelo de pronósticos está basado en el **Modelo Lineal Dinámico (DLM)** de Harrison & Stevens (1976).

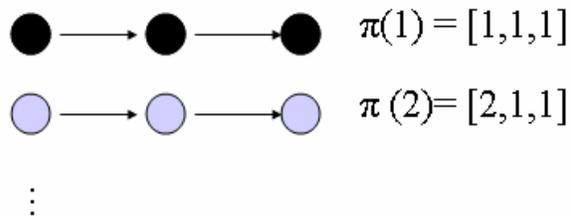
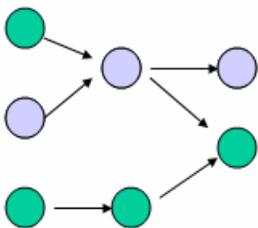
Utiliza una descomposición de la red en subgráficas unilateralmente conectadas con un nodo por clase.



Modelo Lineal Dinámico para Flujos en Redes

El modelo de pronósticos está basado en el **Modelo Lineal Dinámico (DLM)** de Harrison & Stevens (1976).

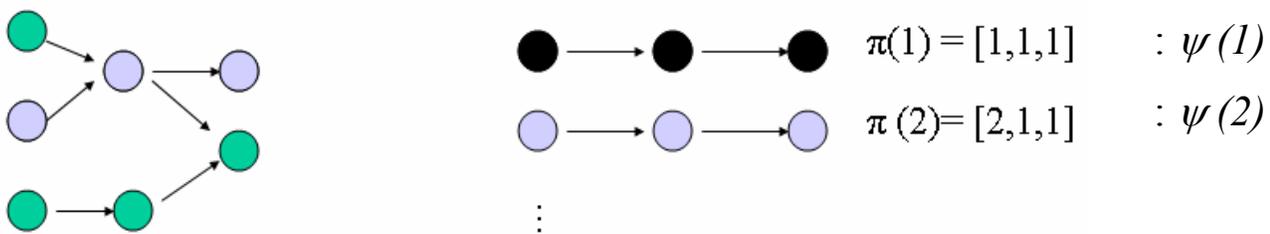
Utiliza una descomposición de la red en subgráficas unilateralmente conectadas con un nodo por clase.



Modelo Lineal Dinámico para Flujos en Redes

El modelo de pronósticos está basado en el **Modelo Lineal Dinámico (DLM)** de Harrison & Stevens (1976).

Utiliza una descomposición de la red en subgráficas unilateralmente conectadas con un nodo por clase.



Modelamos cada flujo ψ sobre las subgráficas y agregamos sobre distintas subgráficas para recuperar los flujos Φ de la red:

$$\Phi_{t+1}(2,1) = \psi_t(1) + \psi_t(2) + \psi_t(3) + \psi_t(4)$$

En general, mantenemos la conservación de la materia

$$\mathbf{1}'\phi_t[l] = \mathbf{1}'\phi_{t+1}[l+1], t \geq l, 1 \leq l < m$$

a través de las ecuaciones “de superposición”

$$\phi_t[l] = \mathbf{S}[l]\psi_{t-l+1}, t \geq l, 1 \leq l \leq m$$

Modelo Lineal Dinámico para Flujos en Redes

RFN DLM

ecuaciones de observación	$\mathbf{X}_t[l] = \phi_t[l] + \nu_{1,t}[l];$	$\nu_{1,t}[l] \sim \mathbf{N}(\mathbf{0}, \mathbf{V}_{1,t}[l])$
	$\mathbf{Y}_t = \psi_{t-m+1} + \nu_{2,t};$	$\nu_{2,t} \sim \mathbf{N}(\mathbf{0}, \mathbf{V}_{2,t})$
ecuación de superposición	$\phi^t = \mathbf{S}\psi_t;$	
ecuación estructural	$\psi_t = \mathbf{F}'_t \theta_t + v_t;$	$v_t \sim \mathbf{N}(\mathbf{0}, \mathbf{U}_t)$
ecuación del sistema	$\theta_t = \mathbf{G}_t \theta_{t-1} + \mathbf{w}_t;$	$\mathbf{w}_t \sim \mathbf{N}(\mathbf{0}, \mathbf{W}_t)$
distribución a priori	$(\theta_0 D_0) \sim \mathbf{N}(\mathbf{a}_{0,0}, \mathbf{R}_{0,0})$	

donde $\phi^t = (\phi_t[1], \phi_{t+1}[2], \dots, \phi_{t+m-1}[m])'$,

$\{\nu_{1,t}[l]\}_{t \geq 1}$, $\{\nu_{2,t}\}_{t \geq m}$, $\{v_t\}_{t \geq 1}$, $\{\mathbf{w}_t\}_{t \geq 1}$ son mutua e internamente independientes y las observaciones están definidas para $t \geq l$ y $t \geq m$, respectivamente.

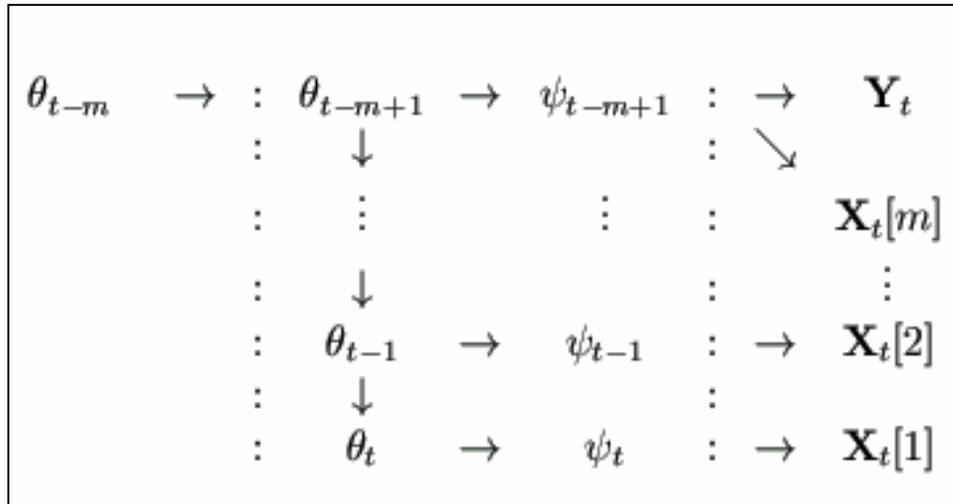
- ✚ cada flujo ψ se modela a través de un Modelo Dinámico Jerárquico (Gamerman & Migon, 1993)
- ✚ combina datos de distintas fuentes, agregados a distintos niveles, secuencialmente de acuerdo a su arribo
- ✚ modela no sólo los flujos en la red, sino su evolución

Las recurrencias son complicadas, pero podemos definir un modelo en el que el parámetro se comporte como en el DLM.

Modelo Lineal Dinámico para Flujos en Redes

Podemos definir un nuevo modelo basado en la propiedad de independencia condicional

$$\{\mathbf{Z}_s\}_{s \geq t} \perp\!\!\!\perp \{\mathbf{Z}_k\}_{k \leq t-1} \mid \{\eta_{t-m+1}, \dots, \eta_t\}; \eta_t = (\theta'_t, \psi'_t)', t > 0, \eta_0 = \theta_0$$



Red Bayesiana para RFN DLM

$$\begin{aligned} \mathbf{Z}_t &= \mathbf{L}'_t \beta_t + \nu_t; & \nu_t &\sim \mathbf{N}(\mathbf{0}, \mathbf{V}_t) \\ \beta_t &= \mathbf{K}_t \beta_{t-1} + \varepsilon_t & \varepsilon_t &\sim \mathbf{N}(\mathbf{0}, \Sigma_t) \\ \beta_0 &\sim \mathbf{N}(\mathbf{m}_0, \mathbf{C}_0) \end{aligned}$$

a) distribución a priori al tiempo $t - 1$, $(\beta_t \mid D_{t-1}) \sim (\mathbf{a}_t, \mathbf{R}_t)$

$$\mathbf{a}_t = \mathbf{K}_t \mathbf{m}_{t-1}, \mathbf{R}_t = \mathbf{K}_t \mathbf{C}_{t-1} \mathbf{K}'_t + \Sigma_t$$

b) distribución predictiva, $(\mathbf{Z}_t \mid D_{t-1}) \sim (\mathbf{f}_t, \mathbf{Q}_t)$

$$\mathbf{f}_t = \mathbf{L}' \mathbf{a}_t, \mathbf{Q}_t = \mathbf{L}' \mathbf{R}_t \mathbf{L} + \mathbf{V}_t$$

c) distribución a posteriori al tiempo t , $(\beta_t \mid D_t) \sim (\mathbf{m}_t, \mathbf{C}_t)$

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t \mathbf{e}_t, \mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t \mathbf{Q}_t \mathbf{A}'_t$$

donde

$$\mathbf{e}_t = \mathbf{z}_t - \mathbf{f}_t, \mathbf{A}_t = \mathbf{R}_t \mathbf{L}_t \mathbf{Q}_t^{-1}$$

$$\mathbf{L}'_t = \begin{pmatrix} & \mathbf{S}[1] & \mathbf{0} & \dots & \mathbf{0} \\ & \mathbf{0} & \mathbf{S}[2] & \dots & \mathbf{0} \\ \mathbf{0}_{(r \sum_{i=1}^t m_i \times tr)} & & \mathbf{0} & \dots & \mathbf{S}[t] \end{pmatrix}; \nu_t = \begin{pmatrix} \nu_{1,t}[1] \\ \nu_{1,t}[2] \\ \vdots \\ \nu_{1,t}[t] \end{pmatrix}$$

si $t < m$, y si $t \geq m$

$$\mathbf{L}'_t = \begin{pmatrix} & \mathbf{S}[1] & \mathbf{0} & \dots & \mathbf{0} \\ & \mathbf{0} & \mathbf{S}[2] & \dots & \mathbf{0} \\ \mathbf{0}_{(n+\rho) \times mr} & & \mathbf{0} & \dots & \mathbf{S}[m] \\ & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_{\rho \times \rho} \end{pmatrix}; \nu_t = \begin{pmatrix} \nu_{1,t}[1] \\ \nu_{1,t}[2] \\ \vdots \\ \nu_{1,t}[m] \\ \nu_{2,t} \end{pmatrix}$$

$$\mathbf{K}_t = \begin{pmatrix} \mathbf{K}_{1,t} & \mathbf{K}_{2,t} \\ \mathbf{K}_{3,t} & \mathbf{K}_{4,t} \end{pmatrix}$$

$$\mathbf{K}_{1,t} = \begin{pmatrix} \mathbf{G}_t & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} \end{pmatrix}_{(tr \times (t-1)r)}; \quad \mathbf{K}_{2,t} = \mathbf{0}_{(tr \times (t-1)\rho)}$$

$$\mathbf{K}_{3,t} = \begin{pmatrix} \mathbf{F}'_t \mathbf{G}_t & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}_{(t\rho \times (t-1)r)}; \quad \mathbf{K}_{4,t} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} \end{pmatrix}_{(t\rho \times (t-1)\rho)}$$

si $t \leq m$, y si $t > m$, entonces

$$\mathbf{K}_{1,t} = \begin{pmatrix} \mathbf{G}_t & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} & \mathbf{0} \\ \vdots & & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \end{pmatrix}_{(mr \times mr)}; \quad \mathbf{K}_{2,t} = \mathbf{0}_{(mr \times m\rho)}$$

$$\mathbf{K}_{3,t} = \begin{pmatrix} \mathbf{F}'_t \mathbf{G}_t & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}_{(m\rho \times m\rho)}; \quad \mathbf{K}_{4,t} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} & \mathbf{0} \\ \vdots & & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \end{pmatrix}_{(m\rho \times m\rho)}$$

y con error del sistema $\varepsilon_t = (\mathbf{w}'_t, \mathbf{0}', \dots, \mathbf{0}', \mathbf{w}'_t \mathbf{F}_t + \nu'_t, \mathbf{0}', \dots, \mathbf{0}')'$, $\varepsilon_0 = \mathbf{w}_0$

Intervención en el Modelo para Flujo en Redes

Manipulaciones típicas a sistemas de flujos son:

- ✚ cierre de calles
- ✚ embargos
- ✚ aumento de flujo

Cálculo de Intervenciones (Spirtes et al., 1993; Pearl, 1995)

Desarrollado para estudiar los efectos de manipulaciones. Se da una **interpretación causal** a un sistema representado por una Red Bayesiana.

Una Red Bayesiana es llamada Causal si la densidad f que respeta la factorización

$$f(x) = \prod_{v \in V} f(x_v | x_{pa_v})$$

respeta bajo la manipulación de las variables en A la factorización

$$f(x_{V \setminus A} || x_A = x_A^*) = \prod_{v \in V \setminus A} f(x_v | x_{pa_v}^*)$$

- ✚ la manipulación borra arcos en la Red Bayesiana
- ✚ otras formas de manipulación exploradas por Pearl y por West & Harrison (1986, 1989) re-asignan padres a los nodos: **modifican las ecuaciones.**

Intervención en el Modelo para Flujo en Redes

Queremos incluir en el modelo nueva información obtenida a partir del conocimiento de la **intención** de algún agente de manipular un proceso del sistema.

En vez de cambiar el valor de una variable o crear/destruir relaciones en el sistema,

- ✚ modificamos las distribuciones de los errores del modelo y calculamos los efectos en otros procesos a través de la **Regla Generalizada de Jeffrey**
- ✚ nos apartamos del marco gráfico para calcular los efectos de manipulaciones más generales: no deseamos comprometernos con una dirección en la relación causal entre Φ y ψ
- ✚ especificamos el tiempo de la manipulación y el tiempo en el que se presentan los efectos

Regla de Jeffrey

(Jeffrey, 1965; Diaconis & Zabell, 1982)

Provee una forma de actualizar nuestro conocimiento cuando no tenemos certeza a posteriori.

Tenemos un audio libro de Shakespeare

O: lector es Olivier

G: lector es Gielgud

a priori $p(O), p(G)$

H: el lector se casó con Vivien Leigh

Queremos saber la probabilidad $p(H)$ y tenemos $p(H|O), p(H|G)$

Escuchamos el audio libro, pero no estamos seguros de quién es el lector.

a posteriori $p^*(O), p^*(G)$

Pero podemos asumir que

$$p^*(H|O)=p(H|O), p^*(H|G)=p(H|G)$$

Calculamos la probabilidad a posteriori $p^*(H)$ por la **Regla de Jeffrey**:

$$p^*(H)=p(H|O) p^*(O)+ p(H|G) p^*(G)$$

Condicionar por la Regla de Bayes es un caso límite de la Regla de Jeffrey.

Intervención en el Modelo para Flujo en Redes

Sean $p(\mathbf{v}_k) = N(\mathbf{a}_k, \mathbf{R}_k)$ y $p(\epsilon_T) = N(\mathbf{a}, \mathbf{R})$ distribuciones marginales dada la información al tiempo $t-1$ para el modelo RFN DLM y sea $p^*(\epsilon_T) = N(\mathbf{a}^*, \mathbf{R}^*)$ la nueva distribución de un error al tiempo $t-1$. Si $\mathcal{A}_0 = \sigma(\epsilon_T)$ es suficiente para la familia $\{p, p^*\}$, entonces la nueva distribución de \mathbf{v}_k al tiempo $t-1$ está dada por $p^*(\mathbf{v}_k) = N(\mathbf{a}_k^*, \mathbf{R}_k^*)$ donde

$$\begin{aligned}\mathbf{a}_k^* &= \mathbf{a}_k + \mathbf{A}_{\mathbf{v}_k, \epsilon_T}(\mathbf{a}^* - \mathbf{a}), \\ \mathbf{R}_k^* &= \mathbf{R}_k + \mathbf{A}_{\mathbf{v}_k, \epsilon_T}(\mathbf{R}^* - \mathbf{R})\mathbf{A}'_{\mathbf{v}_k, \epsilon_T}.\end{aligned}$$

Una vez actualizadas las probabilidades de acuerdo a la manipulación, podemos continuar pronosticando y actualizando probabilidades como de rutina.

- Las variables independientes del error intervenido no tienen un efecto en los tiempos $t, t+1, \dots, T-1$
- Al momento T incluso éstas son afectadas: la distribución retrospectiva cambia

Esta metodología de intervención se puede utilizar para representar manipulaciones en

- observaciones
- flujos por una trayectoria específica
- tendencia, crecimiento, estacionalidad, etc.

Intervención en el Modelo para Flujo en Redes

Al manipular los flujos Φ no deseamos asumir que todas las distribuciones condicionales serán invariantes.

La única relación que se mantiene inmutable es la de la conservación de la materia:

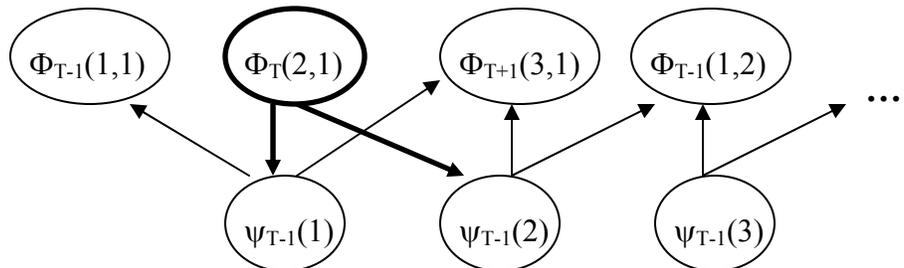
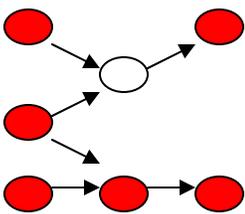
$$\{\psi_{T-l+1}(i); 1 \leq i \leq \rho, (\mathbf{S}[l])_{j_l, i} = 1\}$$

$$p^*(\psi_{T-l+1}^{(l, j_l)} | \phi_T(l, j_l), D_{t-1}) = p(\psi_{T-l+1}^{(l, j_l)} | \phi_T(l, j_l), D_{t-1})$$

Calculamos los efectos en los flujos ψ por las trayectorias que pasan por el nodo intervenido:

$$p^*(\psi_{T-l+1}^{(l, j_l)} | D_{t-1}) = \int p(\psi_{T-l+1}^{(l, j_l)} | \phi_T(l, j_l), D_{t-1}) p^*(\phi_T(l, j_l) | D_{t-1}) d\phi_T(l, j_l)$$

Podemos finalmente transformar esta manipulación en una intervención directa sobre los flujos ψ como en el caso anterior, es decir, manipulando los errores de la ecuación estructural y calcular los efectos en el flujo por otros nodos de la red.



Comentarios Finales

- * Es posible crear un modelo de pronósticos con un cálculo de intervenciones apropiado para redes de flujo y con una **red bayesiana (dinámica)** asociada.
- * El cálculo definido coincide con las teorías ya existentes: intervenir un modelo debido a una manipulación no modifica el pasado, pero puede **modificar nuestras creencias (presentes y futuras) acerca del pasado.**
- * Es posible definir manipulaciones conjuntas.
- * Es necesario extender la teoría de manipulaciones para estudiar el **nacimiento de nuevos procesos** de flujo.
- * Es necesario considerar otras **distribuciones** en los supuestos del modelo: p. ej. Poisson, así como el caso en que las **varianzas son desconocidas.**

Referencias

- [1] Diaconis, P., and Zabell, S.L. (1982). Updating Subjective Probability. *Journal of the American Statistical Association*, 77, 822-830.
- [2] Gamerman, D., and Migon, H. (1993). Dynamic hierarchical models. *J. Roy. Statist. Soc. (Ser. B)* 55, 629-642.
- [3] Harrison, P. J., and Stevens, C. (1976). Bayesian forecasting (with discussion). *J. Roy. Statist. Soc. (Ser. B)*, 38, 205-247.
- [4] Jeffrey, R. C. (1965). *The Logic of Decision*. New York: McGraw-Hill.
- [5] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann.
- [6] Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82, 669-710.
- [7] Spirtes, P., Glymour, C. and Scheines, R. (2000). *Causation, prediction and search*. New York: Springer-Verlag.
- [8] Tebaldi, C., and West, M. (1998). Bayesian inference on network traffic using link count data. *Journal of the American Statistical Association*, 93, 557-576.
- [9] West, M., and Harrison, P.J. (1986). Monitoring and Adaption in Bayesian forecasting models. *J. Amer. Statist. Assoc.*, 81, 741-750.
- [10] West, M. and Harrison, P.J. (1989). Subjective intervention in formal models. *Journal of Forecasting*, 8, 33-53.
- [11] Whitlock, M.E., Queen, C.M. (2000). Modelling a Traffic Network with Missing Data. *Journal of Forecasting*, 19, Issue 7, 561-574.
- [12] Whittaker J., Garside, S., and Lindveld, K. (1997). Tracking and predicting a network traffic process. *International Journal of Forecasting*, 13, 51-61.
- [13] Zohar, R., and Geiger, D. (2002). Estimation in Hidden Flow Models. *Tech. Rep.* Technion, Israel.