Razonamiento Probabilista

Redes de Márkov

Verónica E. Arriola-Rios

Facultad de Ciencias, UNAM

9 de agosto de 2021

Redes de Márkov por pares

Temas

1. Redes de Márkov por pares
   - Definición
   - Flujo de influencia probabilista

Redes de Márkov por pares

Definición

Una *red de Márkov por pares* es una gráfica no dirigida cuyos nodos son las variables aleatorias $X_1, ..., X_n$ y cada arista $X_i - X_j$ se encuentra asociada con un factor (llamado **potencial**) $\phi_{ij}(X_i, X_j)$

- Las redes de Márkov son especialmente útiles para modelar fenómenos donde no se puede atribuir la relación entre dos variables a un principio de causalidad, por lo que no se puede asociar una dirección.
- Los valores asociados en el factor reflejan la *afinidad* entre asignaciones de valores por pares, también llamada *compatibilidad* o *restricciones suaves* (*affinity, compatibility, soft constraints*).

Ejemplo: amigos estudiando

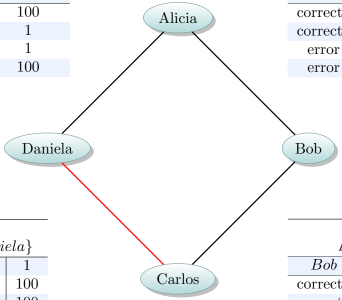
- Los siguientes pares de amigos suelen estudiar juntos, pero algunos tienden a estar de acuerdo, mientras que otros suelen pelear y elegir opiniones opuestas. (Koller 2016)

$\phi_4(D, A)$
$\mathbb{A} = \{Daniela, Alicia\}$

Daniela	*Alicia*	Influencia
correcto	correcto	100
correcto	error	1
error	correcto	1
error	error	100

$\phi_1(A, B)$
$\mathbb{A} = \{Alicia, Bob\}$

Alicia	*Bob*	Influencia
correcto	correcto	30
correcto	error	5
error	correcto	1
error	error	10



$\phi_3(C, D)$
$\mathbb{A} = \{Carlos, Daniela\}$

correcto	correcto	1
correcto	error	100
error	correcto	100
error	error	1

$\phi_2(B, C)$
$\mathbb{A} = \{Bob, Carlos\}$

Bob	*Carlos*	Influencia
correcto	correcto	100
correcto	error	1
error	correcto	1
error	error	100

- La distribución de probabilidad conjunta completa para este sistema es por definición:

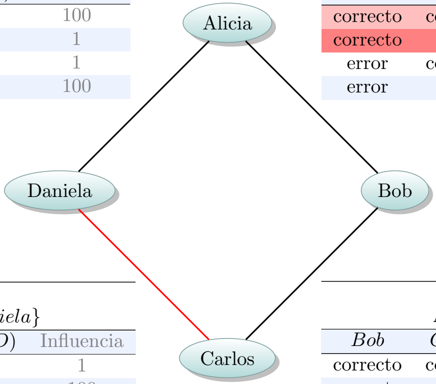
$$\tilde{P}(A,B,C,D) = \phi_1(A,B) \times \phi_2(B,C) \times \phi_3(C,D) \times \phi_4(A,D)$$

$$P = \frac{\tilde{P}(A,B,C,D)}{\sum_{A,B,C,D} \tilde{P}(A,B,C,D)} = \frac{1}{Z}\tilde{P}(A,B,C,D)$$

- A partir de esta tabla es posible obtener cualquier otra distribución de probabilidad mediante las operaciones de marginalización, reducción y normalización.

$P(D, A)$ $\mathbb{A} = \{Daniela, Alicia\}$			
Daniela	*Alicia*	$P(D, A)$	Influencia
correcto	correcto	0.78	100
correcto	error	0.04	1
error	correcto	0.01	1
error	error	0.17	100

$P(A, B)$ $\mathbb{A} = \{Alicia, Bob\}$			
Alicia	*Bob*	$P(A, B)$	Influencia
correcto	correcto	0.12	30
correcto	error	0.69	5
error	correcto	0.14	1
error	error	0.04	10



$P(C, D)$ $\mathbb{A} = \{Carlos, Daniela\}$			
Carlos	*Daniela*	$P(C, D)$	Influencia
correcto	correcto	0.04	1
correcto	error	0.19	100
error	correcto	0.75	100
error	error	0.01	1

$P(B, C)$ $\mathbb{A} = \{Bob, Carlos\}$			
Bob	*Carlos*	$P(B, C)$	Influencia
correcto	correcto	0.22	100
correcto	error	0.04	1
error	correcto	0.01	1
error	error	0.72	100

Temas

1. Redes de Márkov por pares
   - Definición
   - Flujo de influencia probabilista

Ruta activa

Definición (Ruta activa)

Una *ruta* $X_1 - \ldots - X_n$ se encuentra *activa* dado el conjunto de variables observadas Z si ninguna X_i se encuentra en Z.

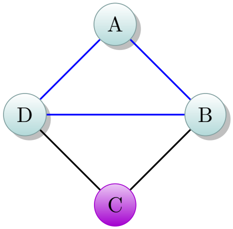


Figura: Habiendo sido observada C, todos los caminos que pasan por ahí están **inactivos**.

Distribución de Gibbs General

1. Redes de Márkov por pares
2. Distribución de Gibbs General

Temas

2. Distribución de Gibbs General
   - Factores generales
   - Distribución de Gibbs
   - Red de Márkov inducida

Factores generales

- Los *factores generales* $\phi_i(\vec{D}_i)$ pueden describir las interacciones entre un subconjunto de k variables aleatorias en $\vec{D}$ (el alcance del factor ϕ_i).
- Estos factores **inducen** una red de Márkov H_Φ donde se agrega una arista $X_i - X_j$ por cada par (X_i, X_j) tal que $X_i, X_j \in \vec{D}_i$.

Temas

2. Distribución de Gibbs General
   - Factores generales
   - Distribución de Gibbs
   - Red de Márkov inducida

Distribución de Gibbs

- Dado el conjunto de factores:

$$\Phi = \{\phi_1(\vec{D}_1), ..., \phi_k(\vec{D}_k)\}$$

- Sea la medida no normalizada

$$\tilde{P}_\Phi(X_1, ..., X_n) = \prod_{i=1}^{k} \phi_i(\vec{D}_i)$$

- Y la función de partición Z, la constante de normalización:

$$Z_\Phi = \sum_{X_1,...,X_n} \tilde{P}_\Phi(X_1, ..., X_n)$$

- La *distribución de Gibbs* representa una distribución de probabilidad como el producto de factores normalizado:

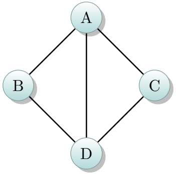
$$P_\Phi(X_1, ..., X_n) = \frac{1}{Z_\Phi} \tilde{P}_\Phi(X_1, ..., X_n) \tag{1}$$

Temas

2. Distribución de Gibbs General
   - Factores generales
   - Distribución de Gibbs
   - Red de Márkov inducida

Red de Márkov inducida

- En general, existen varios conjuntos de factores generales que pueden inducir la misma red de Márkov. Ejemplo:
  - $\phi_1(A, B, D), \phi_2(A, C, D)$
  - $\phi_1(A, B), \phi_2(B, D), \phi_3(D, C), \phi_4(C, A), \phi_5(A, D)$
  - $\phi_1(A, B, D), \phi_2(A, C), \phi_3(C, D)$



  Ojo, esta otra agregaría una arista más:
  - $\phi_1(A, B, D, C), \phi_2(A, D)$
- La red inducida define unívocamente el comportamiento del flujo de influencia probabilista.
- Queda abierto el problema de definir, para un problema dado, cuál es el conjunto de factores que describirá correctamente el comportamiento del sistema modelado con la red de Márkov, es decir, cuál **factoriza** la distribución P.

Referencias I

- Koller, Daphne (15 de jun. de 2016). *Programa especializado: Probabilistic Graphical Models*. Ed. por Coursera. URL: https://www.coursera.org/specializations/probabilistic-graphical-models.
- Koller, Daphne y Nir Friedman (2009). *Probabilistic Graphical Models, Principles and Techniques*. MIT Press Cambridge.

Licencia