

Inferencia en Redes Causales Probabilistas mediante Precomputación Aproximada.

Luis D. Hernández¹, Serafín Moral², Antonio Salmerón³.

¹ Departamento de Informática y Sistemas
Universidad de Murcia

² Departamento de Ciencias de la Computación e I.A.
Universidad de Granada

³ Departamento de Estadística y Matemática Aplicada
Universidad de Almería

RESUMEN

Presentamos una clase de algoritmos de Monte Carlo para la propagación de probabilidades en redes causales. En una primera etapa se calculan las funciones de muestreo de forma aproximada, que luego se utilizan para estimar la distribución a posteriori de las variables de la red.

Palabras y frases clave: Redes causales, muestreo por importancia, precomputación aproximada.

Clasificación AMS: 68T30.

1 Introducción

En respuesta a los problemas que presentan los algoritmos exactos para la inferencia en redes causales de tamaño grande, surgen los métodos basados en técnicas de Monte Carlo, destacando los llamados de muestreo por importancia. En este caso, el punto clave es la obtención de funciones de muestreo tan próximas a las originales como sea posible. En una red causal, la distribución original se presenta como el producto de las distribuciones condicionadas de cada variable a sus padres, por lo que los algoritmos clásicos usan funciones de muestreo próximas a las condicionadas para obtener pesos más uniformes en el proceso de simulación.

En este trabajo proponemos la idea de usar, en cada momento, toda la información disponible para cada variable. Es decir, a la hora de simular una variable, utilizar todas las funciones que están definidas para dicha variable. En el peor de los casos, esto sería tan complejo como el cálculo exacto de las probabilidades. Sin embargo, definiendo criterios para la combinación de funciones, esta operación puede hacerse de forma aproximada, agilizando los cálculos.

2 El algoritmo principal

Para una red de n variables $\{X_1, \dots, X_n\}$, el algoritmo empieza con una familia de funciones dadas por las condicionadas más las observaciones: $H = \{f_1, \dots, f_n\} \cup \{\delta_{e_i}\}_{i \in E}$ donde E es el conjunto de las variables observadas. Se establece un orden de eliminación de las variables σ . Se eliminan en secuencia las variables según el orden σ mediante el siguiente procedimiento:

- Sea $H(i)$ el conjunto de las funciones definidas para la variable $X_{\sigma(i)}$. Eliminar $H(i)$ de H .
- Repetir varias veces el siguiente proceso: Tomar $R \subset H(i)$; combinar todas las funciones en R ; añadir el resultado de la combinación a $H(i)$. Eliminar R de $H(i)$.
- Calcular $H^+(i)$ a partir de $H(i)$ eliminando $X_{\sigma(i)}$ de todas las funciones contenidas en $H(i)$. Añadir $H^+(i)$ a H .

Obsérvese que después de eliminar una variable $X_{\sigma(i)}$, en $H(i)$ tenemos un conjunto de funciones que dependen exclusivamente de $X_{\sigma(i)}$ y otras variables que aún no han sido simuladas. Por lo tanto, cuando lleguemos a la última variable, $X_{\sigma(n)}$, obtendremos una función que depende sólo de dicha variable. Entonces podemos simular un valor para la misma e instanciar las funciones de muestreo de la variable inmediatamente anterior, simulándola después. Repitiendo este proceso para todas las variables, se obtiene una configuración para las n variables de la red. A cada configuración se le asigna un peso o importancia igual al cociente entre la probabilidad exacta de la configuración y su probabilidad de muestreo. Una propiedad importante de la eliminación de las variables es que no se introducen valores cero en este proceso. Además, se puede detectar cuándo se pueden hacer los cálculos exactos, a saber, cuando $R = H(i)$.

Supongamos que utilizamos el algoritmo para obtener valores de las n variables. Para obtener un valor para $X_{\sigma(i)}$, el proceso es el siguiente:

- Sea $H(i)$. Restringir cada función contenida en $H(i)$ a los valores obtenidos para las variables ya simuladas. Combinar todas las funciones contenidas en $H(i)$. Se obtiene entonces una función h'_i definida solamente para $X_{\sigma(i)}$.
- Si $N(h'_i)$ es la normalización de h'_i , obtener un valor para $X_{\sigma(i)}$ siguiendo la distribución de probabilidad $N(h'_i)$.

Referencias

- Hernández, L.D., Moral, S., Salmerón, A. (1996). *A Monte Carlo algorithm for probabilistic propagation based on importance sampling and stratified simulation techniques*. Sometido a: International Journal of Approximate Reasoning.
- Salmerón, A. (1996) *Propagación de probabilidades mediante precomputación aproximada*. Technical Report #GAD9603. Grupo de Análisis de Datos. Universidad de Almería.