知識工学(第 14 回)

二宮 崇 (ninomiya@cs.ehime-u.ac.jp)

確率推論 (14 章)

§14.4.2 変数消去アルゴリズム

同じ計算を二度しないように工夫する動的計画法の一種

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

式の注釈付けしたパートを因子(factor)と呼ぶ。因子は確率分布の表として表される。

A	$f_4(A)\ (= P(j|A))$
0	0.05
1	0.90

A	$f_5(A)\ (= P(m|A))$
0	0.01
1	0.70

$$P(B|j,m) = \alpha f_1(B) \times \sum_E f_2(E) \times \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A)$$

$$\begin{aligned} f_6(B,E) &= \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A) \\ &= f_3(a,B,E) \times f_4(a) \times f_5(a) + f_3(\neg a,B,E) \times f_4(\neg a) \times f_5(\neg a) \end{aligned}$$

$$P(B|j,m) = \alpha f_1(B) \times \sum_E f_2(E) \times f_6(B,E)$$

$$
\begin{aligned}
f_7(B) &= \sum_E f_2(E) \times f_6(B,E) \\
&= f_2(e) \times f_6(B,e) + f_2(\neg e) \times f_6(B,\neg e)
\end{aligned}
$$

$$P(B|j,m) = \alpha f_1(B) \times f_7(B)$$

○因子に対する操作

$$f\left(X_1 \cdots X_j, Y_1 \cdots Y_k, Z_1 \cdots Z_l\right) = f_1\left(X_1 \cdots X_j, Y_1 \cdots Y_k\right) \times f_2(Y_1 \cdots Y_k, Z_1 \cdots Z_l)$$

A	B	$f_1(A,B)$	B	C	$f_2(B,C)$	A	B	C	$f_3(A,B,C)$
0	0	0.1	0	0	0.4	0	0	0	$0.1 \times 0.4 = 0.04$
0	1	0.9	0	1	0.6	0	0	1	$0.1 \times 0.6 = 0.06$
1	0	0.7	1	0	0.8	0	1	0	$0.9 \times 0.8 = 0.72$
1	1	0.3	1	1	0.2	0	1	1	$0.9 \times 0.2 = 0.18$
						1	0	0	$0.7 \times 0.4 = 0.28$
						1	0	1	$0.7 \times 0.6 = 0.42$
						1	1	0	$0.3 \times 0.8 = 0.24$
						1	1	1	$0.3 \times 0.2 = 0.06$

$$f(B,C) = \sum_A f_3(A,B,C) = f_3(\neg a,B,C) + f_3(a,B,C)$$

B	C	$f_3(\neg a,B,C)$	$f_3(a,B,C)$	$f(B,C)$
0	0	0.04	0.28	0.32
0	1	0.06	0.42	0.48
1	0	0.72	0.24	0.96
1	1	0.18	0.06	0.24

○防犯アラームの具体的な解き方

従って、防犯アラームの例は具体的には次のようにして解くことができる。

因子を次のようにおく。

$$P(B|j,m) = \alpha \underbrace{P(B)}_{f_1(B)} \sum_E \underbrace{P(E)}_{f_2(E)} \sum_A \underbrace{P(A|B,E)}_{f_3(A,B,E)} \underbrace{P(j|A)}_{f_4(A)} \underbrace{P(m|A)}_{f_5(A)}$$

B	$f_1(B)\ \big(= P(B)\big)$
0	0.999
1	0.001

E	$f_2(E)\ \big(= P(E)\big)$
0	0.998
1	0.002

A	B	E	$f_3(A,B,E)\ \big(= P(A|B,E)\big)$
0	0	0	0.999
0	0	1	0.71
0	1	0	0.06
0	1	1	0.05
1	0	0	0.001
1	0	1	0.29
1	1	0	0.94
1	1	1	0.95

A	$f_4(A)\ \big(= P(j|A)\big)$
0	0.05
1	0.90

A	$f_5(A)\ \big(= P(m|A)\big)$
0	0.01
1	0.70

A	B	E	$f_3(A,B,E)$	$f_4(A)$	$f_5(A)$	$f_3(A,B,E) \times f_4(A) \times f_5(A)$
0	0	0	0.999	0.05	0.01	0.0004995
0	0	1	0.71	0.05	0.01	0.000355
0	1	0	0.06	0.05	0.01	0.00003
0	1	1	0.05	0.05	0.01	0.000025
1	0	0	0.001	0.90	0.70	0.00063
1	0	1	0.29	0.90	0.70	0.1827
1	1	0	0.94	0.90	0.70	0.5922
1	1	1	0.95	0.90	0.70	0.5985

$$\begin{aligned} f_6(B,E) &= \sum_A f_3(A,B,E) \times f_4(A) \times f_5(A) \\ &= f_3(a,B,E) \times f_4(a) \times f_5(a) + f_3(\neg a,B,E) \times f_4(\neg a) \times f_5(\neg a) \end{aligned}$$

B	E	$f_6(B,E)$
0	0	$0.0004995 + 0.00063 = 0.0011295$
0	1	$0.000355 + 0.1827 = 0.183055$
1	0	$0.00003 + 0.5922 = 0.59223$
1	1	$0.000025 + 0.5985 = 0.598525$

B	E	$f_2(E)$	$f_6(B,E)$	$f_2(E) \times f_6(B,E)$
0	0	0.998	0.0011295	0.001127241
0	1	0.002	0.183055	0.00036611
1	0	0.998	0.59223	0.59104554
1	1	0.002	0.598525	0.00119705

$$\begin{aligned} f_7(B) &= \sum_E f_2(E) \times f_6(B,E) \\ &= f_2(e) \times f_6(B,e) + f_2(\neg e) \times f_6(B,\neg e) \end{aligned}$$

B	$f_7(B)$
0	$0.001127241 + 0.00036611 = 0.001493351$
1	$0.59104554 + 0.00119705 = 0.59224259$

$$P(B|j,m) \quad = \alpha f_1(B) \times f_7(B)$$

B	$f_1(B)$	$f_7(B)$	$f_1(B) \times f_7(B)$	$P(B|j,m)$
0	0.999	0.001493351	0.001491857649	0.715828
1	0.001	0.59224259	0.00059224259	0.284172

§14.4.3 厳密推論の計算量

多重木(polytree): 任意の二つのノード間に多くても一つの無向パスしか存在しないネットワーク。単結合ネットワーク。

多重木の計算量: CPT の項目総数に対し線形。親の数が定数で制限→ノード数に対し線形。

複結合ネットワーク: 向きを考慮しない場合にループが存在するネットワーク

複結合ネットワークの計算量: 指数的な時間・空間的計算量。#P 困難。

§14.5 ベイジアンネットの近似推論

基本的にベイジアンネットの厳密推論には、ノード数に対し指数的な時間を必要とする。また、動的計画法などの実装も簡単ではない。そこで、近似的に推論をする方法を導入する。ここでは、ランダムサンプリングアルゴリズム(またはモンテカルロアルゴリズム)と呼ばれる方法を紹介する

§14.5.1 ダイレクトサンプリング法

〇棄却サンプリング

0から1の間の乱数が生成されるとする。すると、この乱数を用いて、ベイジアンネットの確率分布に従ったサンプル(全確率変数に対する具体的な値)を生成することができる。推論式$P(X|e_1,\cdots,e_n)$を計算するには、サンプリングを繰り返して、次の計算を行えば良い。

$$P(X|e_1,\cdots,e_n) \approx \frac{N(X,e_1,\cdots,e_n)}{N(e_1,\cdots,e_n)}$$

ただし、全ての変数集合を$\{X\} \cup E_1,\cdots,E_n \cup Y_1,\cdots,Y_m$としたとき、$N(e_1,\cdots,e_n)$は$E_1,\cdots,E_n = e_1,\cdots,e_n$となったサンプルの数であり、$N(X,e_1,\cdots,e_n)$は$E_1,\cdots,E_n = e_1,\cdots,e_n$かつ$X$となったサンプルの数である。

サンプリングはベイジアンネットの根ノードからスタートし、各ノード毎に乱数を発生させ、CPT に従って、各ノードの値を決定していく。根ノードからスタートして、子のノードの値を順に決定していくため、きちんとした順序で決定していけば親のノードが全て決定した後に子ノードの値を決定することができる。例えば、次の CPT があったとき、

B	E	$P(\neg a)$	$P(a)$
0	0	0.999	0.001
0	1	0.71	0.29
1	0	0.06	0.94
1	1	0.05	0.95

親の値が$B=0,E=1$、乱数をrとしたとき、$r<0.71$ならば$A=0$とし、$r \geq 0.71$ならば$A=1$とすれば良い。(連続確率変数の場合は、累積分布関数の逆関数から求める)

〇尤度重み付け法

棄却サンプリングは単純明快でわかりやすい方法であるが、$E_1,\cdots,E_n = e_1,\cdots,e_n$とならなかったサンプルは計算にまったく使われないため、無駄の多いサンプリングといえる。特に$e_1,\cdots,e_n$が非常に稀な現象であるとき、$E_1,\cdots,E_n = e_1,\cdots,e_n$となるサンプルがなかなか得られない。そこで、$E_1,\cdots,E_n = e_1,\cdots,e_n$となるように強制してサンプリングする方法が望ましい。ただし、そのようにした場合は本来の分布に従ったサンプルが得られないため$E_1,\cdots,E_n = e_1,\cdots,e_n$とすることによる補正(重み付け)が必要となる。そのように重み付けを加えることを尤度重み付けという。

今、サンプルを一つ生成しようとしているとする。wを 1.0とする。棄却サンプリングと同様の方法で各ノードの値を順次決定していくが、証拠eに対応する確率変数Eのノードを決定するときは、乱数を生成せず、$E=e$とする。ただし、CPT により得られる$P(E=e|parents(E))$に対し、$w \leftarrow w \times P(E=e|parents(E))$とする。この手順を繰り返し、サンプル一つ生成したところで得られるwがそのサンプルに対する重みとなる。

棄却サンプリングではサンプル一つを 1 個と数えたが、これをw個と数えれば尤度重み付けによるサンプリングとなる。

サンプリングの手法としては、マルコフ連鎖モンテカルロ (Markov chain Monte Carlo: MCMC) が有名であり、その一種であるギブスサンプラー (Gibbs sampler) や、メトロポリス-ヘイスティングサンプラー (Metropolis-Hastings sampler)がよく用いられている。