АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

Сергей Николенко

СПбГУ — Санкт-Петербург

26 ноября 2020 г.

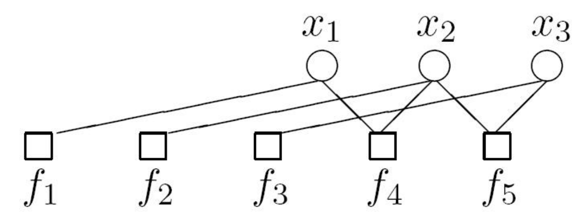
Random facts:

- 26 ноября 1812 г. началось сражение на Березине, в итоге которого «Великая армия» Наполеона была полностью уничтожена; а 26 ноября 1835 г. в ходе Техасской революции произошла Битва за сено, в которой были убиты три мексиканца и ранены четверо техасцев
- 26 ноября 1924 г. Великий народный хурал провозгласил Монгольскую Народную Республику и утвердил конституцию второго в мире социалистического государства
- 26 ноября 1939 г. произошёл Майнильский инцидент, советская провокация, в результате которой началась советско-финская война
- 26 ноября 1941 г. ударное соединение под командованием вице-адмирала Тюити Нагумо покинуло базу в заливе Хитокаппу (Касатка) на Курильских островах и направилось к Пёрл-Харбору

Графические вероятностные модели

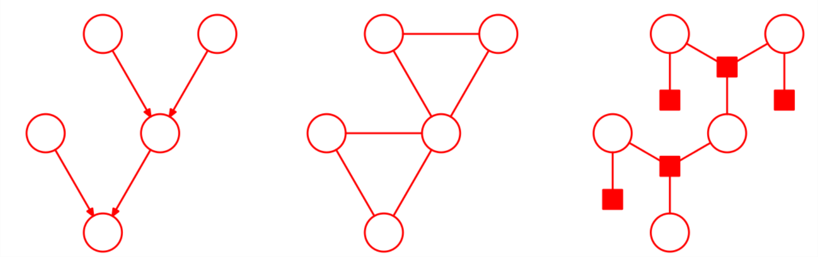
Фактор-графы

- Важная для вывода модификация – *фактор-граф* (можно построить и по направленной модели, и по ненаправленной).
- Фактор-граф – двудольный граф функций и переменных.
- Функция, соответствующая графу, – произведение всех входящих в него функций (т.е. то самое разложение и есть).
- Пример: $p(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_1, x_2) f_5(x_2, x_3)$.



Алгоритм передачи сообщений

Три представления



Функция в общем виде

- Чтобы поставить задачу в общем виде, рассмотрим функцию

$$p^*(X) = \prod_{j=1}^{m} f_j(X_j),$$

  где $X = \{x_i\}_{i=1}^{n}$, $X_j \subseteq X$.
- Т.е. мы рассматриваем функцию, которая раскладывается в произведение нескольких других функций.

Задачи

- Задача нормализации: найти $Z = \sum_X \prod_{j=1}^m f_j(X_j)$.
- Задача маргинализации: найти

$$p_i^*(x_i) = \sum_{k \neq i} p^*(X).$$

  Также может понадобиться, например, $p_{i_1 i_2}$, но реже.
- Поиск гипотезы максимального правдоподобия:

$$\mathbf{x}^* = \arg\max_X p(X).$$

Задачи

- Все эти задачи NP-трудные.
- То есть, если мир не рухнет, сложность их решения в худшем случае возрастает экспоненциально.
- Но можно решить некоторые частные случаи.

Пример

- Давайте начнём с графа в виде (ненаправленной) цепи:

$$p(x_1, \dots, x_n) = \frac{1}{Z}\psi_{1,2}(x_1, x_2) \dots \psi_{n-1,n}(x_{n-1}, x_n).$$

- Мы хотим найти

$$p(x_k) = \sum_{x_1} \dots \sum_{x_{k-1}} \sum_{x_{k+1}} \dots \sum_{x_n} p(x_1, \dots, x_n).$$

Пример

- Очевидно, тут можно много чего упростить; например, справа налево:

$$\sum_{x_n} p(x_1, \dots, x_n) =$$
$$= \frac{1}{Z} \psi_{1,2}(x_1, x_2) \dots \psi_{n-2,n-1}(x_{n-2}, x_{n-1}) \sum_{x_n} \psi_{n-1,n}(x_{n-1}, x_n).$$

- Эту сумму можно вычислить отдельно и продолжать в том же духе справа налево, потом аналогично слева направо.

Пример

- В итоге процесс сойдётся на узле x_k, куда придут два «сообщения»: слева

$$\mu_\alpha(x_k) = \sum_{x_{k-1}} \psi_{k-1,k}(x_{k-1}, x_k) \left[\dots \sum_{x_2} \psi_{2,3}(x_2, x_3) \left[\sum_{x_1} \psi_{1,2}(x_1, x_2) \right] \dots \right],$$

справа

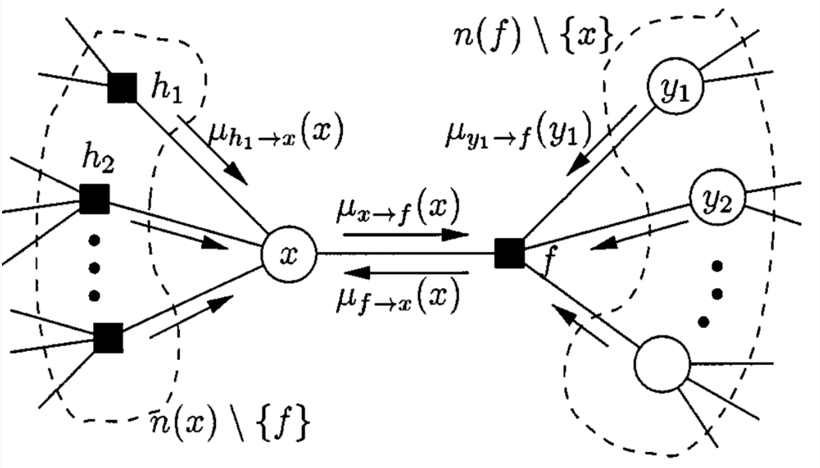
$$\mu_\beta(x_k) = \sum_{x_{k+1}} \psi_{k,k+1}(x_k, x_{k+1}) \left[\dots \left[\sum_{x_n} \psi_{n-1,n}(x_{n-1}, x_n) \right] \dots \right].$$

- Каждую частичную сумму можно рассматривать как «сообщение» от узла к своему соседу, причём это сообщение – функция от соседа.

Алгоритм передачи сообщений

- Чтобы обобщить, удобно рассмотреть опять фактор-граф.
- Предположим, что фактор-граф – дерево (если не дерево, так просто не сработает).
- Алгоритм передачи сообщений решает задачу маргинализации для функции вида $p(x_1, \dots, x_n) = \prod_s f_s(X_s)$, заданной в виде фактор-графа.
- Передаём сообщения по направлению к нужному узлу от переменных к функциям и наоборот.

Передача сообщений



Алгоритм передачи сообщений

- Чтобы найти $p(x_k)$, запишем $p(x_1, \ldots, x_n) = \prod_{s \in \neq(x_k)} F_s(x_k, X_s)$, где X_s – переменные из поддерева с корнем в f_s. Тогда

$$p(x_k) = \sum_{x_{i \neq k}} p(x_1, \ldots, x_n) = \prod_{s \in \neq(x_k)} \left[\sum_{X_s} F_s(x_k, X_s) \right] = \\ = \prod_{s \in \neq(x_k)} \mu_{f_s \to x_k}(x_k),$$

где $\mu_{f_s \to x_k}(x_k)$ – сообщения от соседних функций к переменной x_k.

- Чтобы найти $\mu_{f_s \to x_k}(x_k)$, заметим, что $F_s(x_k, X_s)$ тоже можно разложить по соответствующему подграфу:

$$F_s(x_k, X_s) = f_s(x_k, Y_s) \prod_{y \in Y_s} G_y(y, X_{s,y}),$$

  где Y_s – переменные, непосредственно связанные с f_s (кроме x_k), $X_{s,y}$ – соответствующие поддеревья.
- Итого получаем

$$\mu_{f_s \to x_k}(x_k) = \sum_{Y_s} f_s(x_k, Y_s) \prod_{y \in Y_s} \left(\sum_{X_{s,y}} G_y(y, X_{s,y}) \right) =$$
$$= \sum_{Y_s} f_s(x_k, Y_s) \prod_{y \in Y_s} \mu_{y \to f_s}(y).$$

Алгоритм передачи сообщений

- Можно аналогично подсчитать, что $\mu_{y \to f_s}(y) = \prod_{f \in \neq(y)\ f_s} \mu_{f \to y}(y)$.

Алгоритм передачи сообщений

- Итак, получился простой и понятный алгоритм:
  - как только узел получил сообщения от всех соседей, кроме одного, он сам начинает передавать сообщение в этого соседа;
  - сообщение по ребру между функцией и переменной является функцией от этой переменной;
  - узел-переменная x передаёт сообщение

$$\mu_{x\to f}(x) = \prod_{g \in \neq(x)\ f} \mu_{g\to x}(x);$$

  - узел-функция $f(x, Y)$ передаёт сообщение

$$\mu_{f\to x}(x) = \sum_{y\in Y} f(x, Y) \prod_{y\in Y} \mu_{y\to f}(y);$$

  - начальные сообщения в листьях $\mu_{x\to f}(x) = 1$, $\mu_{f\to x}(x) = f(x)$.

Алгоритм передачи сообщений

- Когда сообщения придут из всех соседей в какую-то переменную x_k, можно будет подсчитать

$$p(x_k) = \prod_{f \in \neq(x_k)} \mu_{f \to x_k}(x_k).$$

- Когда сообщения придут из всех соседей в какой-то фактор $f_s(X_s)$, можно будет подсчитать совместное распределение

$$p(X_s) = f_s(X_s) \prod_{y \in \neq(f_s)} \mu_{y \to f_s}(y).$$

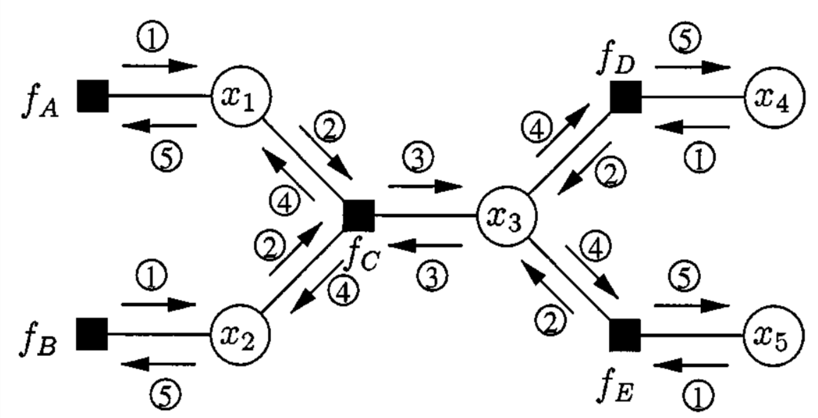
- За два прохода (по каждому ребру туда и обратно) можно будет подсчитать маргиналы во всех узлах.

- Это называется алгоритм sum-product, потому что сообщение вычисляется как

$$\mu_{f\to x}(x) = \sum_{y\in Y} f(x, Y) \prod_{y\in Y} \mu_{y\to f}(y).$$

- Задача максимизации $\arg\max_x p(x_1, \dots, x_n)$ решается так же, но алгоритмом max-sum: сумма заменяется на максимум, а произведение на сумму.

Передача сообщений



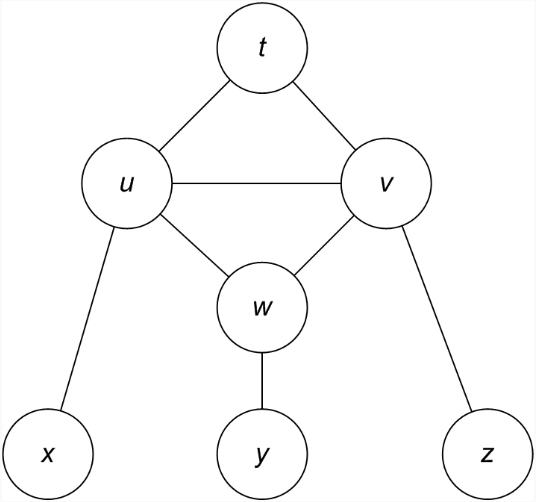
Так что же делать с байесовской сетью?

Для модели не в виде фактор-графа надо просто представить её в виде фактор-графа тем или иным способом.

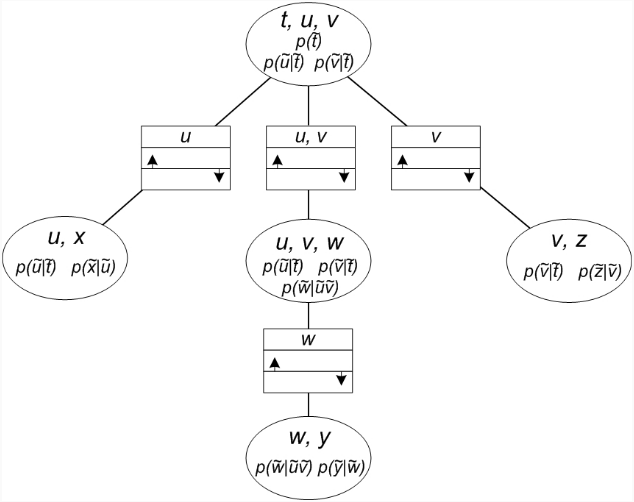
Для байесовской сети это может означать, что надо сначала сделать морализацию, а потом добавить факторы в явном виде.

Так что же делать с байесовской сетью?

Так что же делать с байесовской сетью?



Так что же делать с байесовской сетью?



Спасибо!

Спасибо за внимание!