Заметки к курсу mc201

19 октября 2012 г.

1 Что такое байесовская сеть?

Здесь картинка.
На картинке:

1. Кружочками обозначаются случайные величины.
2. Стрелочками — причинно следственные связи:

   $$X \longrightarrow Y$$

   Значение величины X становится известно раньше значиня Y. Закон распределения величины Y зависит от значения величины X.

Несколько терминов:

1. Вилка (fork)
2. Коллайдер, перевернутая вилка (collider, inverted fork)
3. Путь[1] (trail, path) от A до B — последовательность вершин от вершины A до вершины B, в которой переходы могут делаться и по стрелочкам и против стрелочек
4. Направленный путь от A до B — последовательность вершин от вершины A до вершины B, в которой переходы делаются только по стрелочкам
5. Потомок. Узел Y называется потомком узла X, если существует направленный путь от X до Y.
6. Предок. Узел X называют предком узла Y, если существует направленный путь от X до Y.
7. Прямой родитель. Узел X — прямой родитель узла Y, если $X \to Y$.

По байесовской сети легко определить зависимость и условную зависимость величин. Сначала разберемся с зависимостью.

Определение. Направленный граф без циклов G называется байесовской сетью случайных величин $X_1, \ldots, X_n$, если для любого узла X выполнено условие:

$$X \perp \text{non descendant}(X) \mid \text{parents}(X) \tag{1}$$

Теорема. Величины X и Y независимы, если выполнены все три условия

1. Нет направленного пути от X до Y
2. Нет направленного пути от Y до X
3. Не существует такой величины Z, от которой был бы направленный путь и до X и до Y

Упражнение. Найдите все пары независимых величин.

Нарисовать какую-нибудь картинку

[1] В теории графов термином путь называют то, что мы называем направленным путем

2 Условная независимость

Условная независимость. События A и B называются условно независимыми при условии, что событие C произошло, если $\mathbb{P}(AB \mid C) = \mathbb{P}(A \mid C)\mathbb{P}(B \mid C)$

Примеры.

1. Независимые, но условно зависимые события.
2. Зависимые, но условно независимые события.
3. Независимы при условии C, зависимы при отрицании C

Дискретные случайные величины X и Y условно независимы при условии Z, если для любых x, y и z:

$$\mathbb{P}(X = x, Y = y \mid Z = z) = \mathbb{P}(X = x \mid Z = z) \cdot \mathbb{P}(Y = y \mid Z = z) \tag{2}$$

Условную независимость величин обозначают $X \perp Y \mid Z$

Примечание: некоторые авторы пишут $A \perp B \mid C$ для событий, под этой записью подразумевается на самом деле сразу два условия:

$$A \perp B \mid C \Leftrightarrow \begin{cases} A \text{ и } B \text{ независимы при условии } C \\ A \text{ и } B \text{ независимы при условии } C^c \end{cases} \tag{3}$$

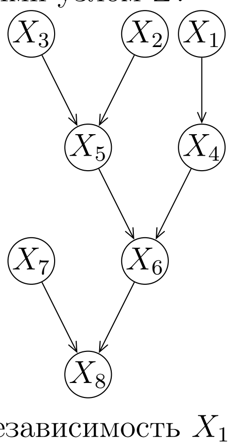
Определение. Путь между X и Y называют d-разделенным (d-separated, directionally separated) множеством узлов Z если выполнено хотя бы одно из условий

1. узел из Z разрывает последовательное соединение на пути
2. узел из Z разрывает «вилку» на пути
3. на пути есть «коллайдер», не являющийся узлом из Z и не содержащий узел из Z в качестве одного из потомков

Можно эквивалентно говорить о том, что путь между X и Y НЕ является d-разделенным узлом Z, если выполнены оба условия:

1. любой коллайдер на пути либо сам является узлом из множества Z, либо имеет потомка из множества Z
2. никакой другой узел на пути не входит в множество Z

Случайные величины X и Y условно независимы при условии Z, если узлы X и Y являются d-разделенными узлом Z.



Упражнения

Проверьте независимость $X_1 \perp X_2$, $X_1 \perp X_2 \mid X_8$, $X_1 \perp X_2 \mid X_7$, $X_1 \perp X_2 \mid X_6$