

# adPredictor and Message Passing

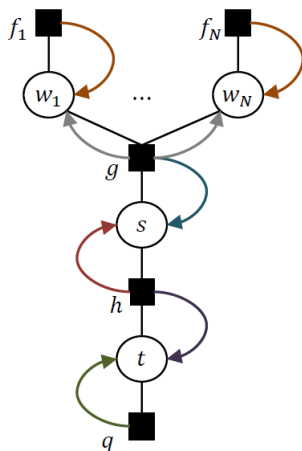
Thore Graepel et al., ICML, 2010

Presented by Boyoung Kim

April 25, 2018

# adPredictor에서 Factor graph

- 논문에서 모형은 잠재변수  $s, t$  를 도입하여 다음과 같은 factor graph로 나타낼 수 있음



# adPredictor에서 Factor graph

- 이때 factor들에 대해 다음과 같은 가정을 하면
  - Factor  $f_i$ : Sample weights  $\mathbf{w}$  from the Gaussian prior  $p(\mathbf{w})$ .
  - Factor  $g$ : Calculate the score  $s$  for  $\mathbf{x}$  as the inner product  $\mathbf{w}^T \mathbf{x}$ , such that  $p(s|\mathbf{x}, \mathbf{w}) := \delta(s = \mathbf{w}^T \mathbf{x})$ .
  - Factor  $h$ : Add zero-mean Gaussian noise to obtain  $t$  from  $s$ , such that  $p(t|s) := \mathcal{N}(t; s, \beta^2)$ .
  - Factor  $q$ : Determine  $y$  by a threshold on the noisy score  $t$  at zero, such that  $p(y|t) := \delta(y = \text{sign}(t))$ .
- 다음과 같이 Joint density를 분해할 수 있음.

$$p(y, t, s, \mathbf{w}|\mathbf{x}) = p(y|t) \cdot p(t|s) \cdot p(s|\mathbf{x}, \mathbf{w}) \cdot p(\mathbf{w})$$

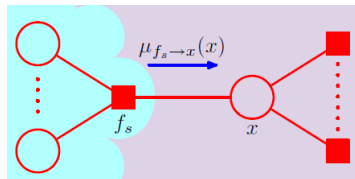
# Message passing algorithm

- Message passing algorithm : factor graph에서 관측된 변수 노드의 분포를 이용하여 관측되지 않은 변수의 주변분포 추론.
- 변수  $x$  의 주변분포를 알고자 한다.

$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x})$$

- 이는  $x$  의 모든 근방의 factor들  $f_s$  로 부터의 message 의 곱으로 계산된다.

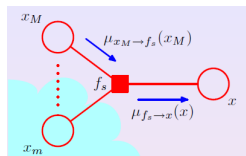
$$p(x) = \prod_{f_s \in ne(x)} \mu_{f_s \rightarrow x}(x)$$



# Message passing algorithm

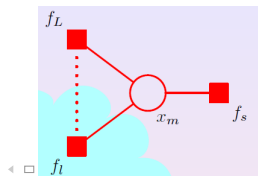
- Message from factor to variable node

$$\mu_{f_s \rightarrow x}(x) = \sum_{x_1} \cdots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{x_m \in ne(f_s) \setminus x} \mu_{x_m \rightarrow f_s}(x_m)$$



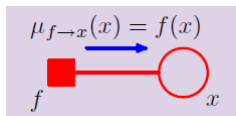
- Message from variable to factor node

$$\mu_{x_m \rightarrow f_s}(x_m) = \prod_{f_l \in ne(x_m) \setminus f_s} \mu_{f_l \rightarrow x_m}(x_m)$$

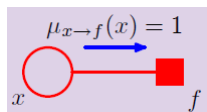


# Message passing algorithm

- 초기값 설정
  - message는 끝노드들(leaves)로부터 (factor가 될수도, variable 이 될수도 있다.) 시작.
  - leaf factor로부터의 메시지는 그 factor 자체로 초기화



- leaf variable로부터의 메시지는 1로 초기화



# Message passing algorithm

- Message passing 체계
  - 분포를 알고자하는 노드를 root라 하자.
  - 끝 노드들 (leaves) 로부터 근방의 노드들로 message가 보내진다.
  - 각각의 내부 노드들이 다른 근방 노드들로부터 message를 받으면 바로 root 쪽으로 message를 보낸다.
  - root가 모든 message를 받으면 root의 주변분포는 이들의 곱으로 계산된다.

- 출처 : Andrea Passerini's slide

- 논문의 사후분포 추론 계산 과정은 Dong Guo's Blog 참고 :

<http://dongguo.me/blog/2013/12/01/bayesian-ctr-prediction-for-bing/>