UCLA Samueli School of Engineering

CS161 WEEK9 DISCUSSION 1C

Danfeng Guo

Agenda

EM ALGORITHM

EM ALGORITHM

It is a method used to find the models that depend on some latent variables

Not only used in unsupervised learning but also supervised learning

It works as long as it has some latent variables

EM ALGORITHM

- Problem setting:

Dataset $Z = \{X, Y\}$ we only know X
The complete log-likelihood is $L(\theta | X, Y)$
θ is the model params we want to know

- E-Step
  - Compute the expected log-likelihood given the observed data and params
- M-Step
  - Compute the params that maximize the log-likelihood

θ_{old}

θ_{new}

EM ALGORITHM

- General concept of EM

$$\log p(X|\theta) = \log \sum_y p(X,y|\theta) = \log \sum_y q(y) \frac{p(X,y|\theta)}{q(y)}$$

$$\geq \sum_y q(y) \log\left(\frac{p(X,y|\theta)}{q(y)}\right)$$

$$\geq \sum q(y) \log\left(\frac{p(y|X,\theta)\,p(X|\theta)}{q(y)}\right)$$

$$\geq \sum q(y) \log\left(\frac{p(y|X,\theta)}{q(y)}\right) + \left(\sum q(y)\right) \log p(X|\theta)$$

$$\geq \left(-KL(q(y), p(y|X,\theta))\right) + \log p(X|\theta)$$

$$L(q(), \theta)$$

KL-Divergence

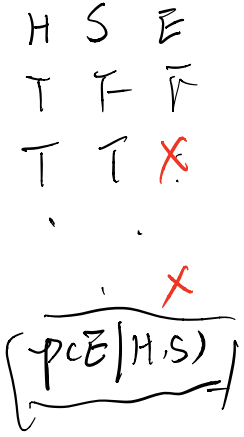
E-Step: Max $L(q,\theta)$ over q

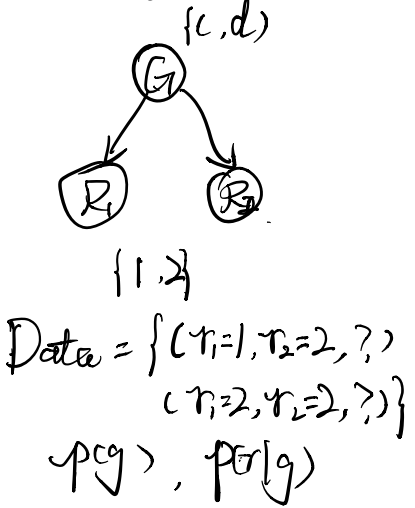
$q(y) = p(y|X,\theta_{old})$

M-Step: Maximize over θ

EM ALGORITHM

- EM in Bayesian Network





Initialize

g	P(g)
c	0.5
d	0.5

g	r	P(r|g)
c	1	0.4
c	2	0.6
d	1	0.6
d	2	0.4

E:

r1,r2	g	P(g,r1,r2)	P(g)
(2,2)	c	0.5×0.6×0.6	0.09
(2,2)	d	0.5×0.4×0.4	0.31
(1,2)	c	0.5×0.4×0.6	0.5
(1,2)	d	0.5×0.6×0.4	0.5

M:

g	P(g)
c	0.09+0.5 ⇒ 0.59
d	0.31+0.5 ⇒ 0.41

g	r	P(r|g)
c	1	0.5
c	2	0.09×2+0.5
d	1	0.5
d	2	0.31×2+0.5

P(r|g)

0.21
0.79
0.31
0.69

Q&A