Artificial Intelligence

Homework 4

Ning Ma

4.1

(a)

$$argmin_{\beta} \sum_{i=1}^{n} |y_i - X_i^T \beta| + \lambda_n \sum_{j=1}^{p} |\beta_j| \tag{1}$$

$$P(C = c|A = a, B = b) = \frac{\sum_{t=1}^{T} I(c, c_t) I(b, b_t) I(a, a_t)}{\sum_{t=1}^{T} I(b, b_t) I(a, a_t)} \tag{2}$$

$$P(D = d|A = a, B = b, C = c) = \frac{\sum_{t=1}^{T} I(d, d_t) I(c, c_t) I(b, b_t) I(a, a_t)}{\sum_{t=1}^{T} I(c, c_t) I(b, b_t) I(a, a_t)} \tag{3}$$

(b)

$$P(a, c|b, d) = \frac{P(a, c, b, d)}{\sum_{a,c} P(a, c, b, d)} \tag{4}$$

$$= \frac{P(a)P(b|a)P(c|b, a)P(d|c, b, a)}{\sum_{a,c} P(a)P(b|a)P(c|b, a)P(d|c, b, a)} \tag{5}$$

(c)

$$P(a|b, d) = \sum_{c} P(a, c|b, d) \tag{6}$$

$$P(c|b, d) = \sum_{a} P(a, c|b, d) \tag{7}$$

(d)

$$L = \sum_t log P(B = b_t, D = d_t) \tag{8}$$
$$= \sum_t log \sum_{a,c} P(a, B = b_t, c, D = d_t) \tag{9}$$
$$= \sum_t log \sum_{a,c} P(a)P(B = b_t|a)P(c|B = b_t, a)P(D = d_t|c, B = b_t, a) \tag{10}$$

(e)

$$P(X_i = x|pa_i = \pi) = \frac{\sum_t P(X_i = x, pa_i = \pi|V = v(t))}{\sum_t P(pa_i = \pi|V = v(t))} \tag{11}$$

$$P(A = a|Pa_A = \phi) = \frac{\sum_t P(A = a|b_t, d_t)}{\sum_t P(Pa_A = \phi|b_t, d_t)} \tag{12}$$
$$= \frac{\sum_t P(a|b_t, d_t)}{\sum_t 1} \tag{13}$$

$$P(B = b|A = a) = \frac{\sum_t P(A = a, B = b|b_t, d_t)}{\sum_t P(A = a|b_t, d_t)} \tag{14}$$
$$= \frac{\sum_t I(b, b_t)P(a|b_t, d_t)}{\sum_t P(a|b_t, d_t)} \tag{15}$$

$$P(C = c|A = a, B = b) = \frac{\sum_t P(C = c, A = a, B = b|b_t, d_t)}{\sum_t P(A = a, B = b|b_t, d_t)} \tag{16}$$
$$= \frac{\sum_t I(b, b_t)P(a, c|b_t, d_t)}{\sum_t I(b, b_t)P(a|b_t, d_t)} \tag{17}$$

$$P(D = c|A = a, B = b, C = c) = \frac{\sum_t P(D = d, A = a, B = b, C = c|b_t, d_t)}{\sum_t P(A = a, B = b, C = c|b_t, d_t)} \tag{18}$$
$$= \frac{\sum_t I(d, d_t)I(b, b_t)P(a, c|b_t, d_t)}{\sum_t I(b, b_t)P(a, c|b_t, d_t)} \tag{19}$$

4.2

$$
\begin{aligned}
L &= \sum_t logP(y_t|\vec{x}_t) && (20)\\
&= \sum_t [(1-y_t)logP(y=0|\vec{x}_t) + y_t logP(y=1|\vec{x}_t)] && (21)\\
&= \sum_t [(1-y_t)log\prod_{i=1}^{n}(1-p_i)^{x_{it}} + y_t log(1-\prod_{i=1}^{n}(1-p_i)^{x_{it}})] && (22)\\
&= \sum_t [(1-y_t)\sum_{i=1}^{n} x_{it} log(1-p_i) + y_t log(1-\prod_{i=1}^{n}(1-p_i)^{x_{it}})] && (23)
\end{aligned}
$$

After 64 iterations I got the following table:

num	0	1	2	4	8	16	32	64
L	-7088.1	-6746.8	-6617.7	-6539.6	-6475.1	-6391.5	-6345.3	-6335.1

The estimates for the parameter p_i are

```
P =

    0.5134
    0.3151
    0.2902
    0.1613
    0.1733
    0.1569
    0.1245
    0.0917
    0.0569
    0.0883
    0.0753
    0.0691
    0.0798
    0.0758
    0.0861
    0.0768
    0.0679
    0.0590
```

The following is the Matlab source code

```
-------------------calculate the Likelihood---------------------

function L=cse_hw4_2_L(X,Y,P,n,T)
%calculate the log likelihoood in each iteration
%n=size(X,2);
%T=size(X,1);
L=0;
for t=1:T
    sum=0;
    product=1;
    for i=1:n
        sum=sum+X(t,i)*log(1-P(i));
    end
    for i=1:n
        product=product*(1-P(i))^(X(t,i));
    end
    L=L+(1-Y(t))*sum+Y(t)*log(1-product);
end

---------------------calculate the updated probability-----------------

function P=cse_hw4_2_P(X,Y,Pb,n,T)
%calculate the updated probability in each iteration
%n is the number of diseases, T is the number of samples(patients),Pb is
%the estimation of probability before update
P=Pb;
for i=1:n
    summation=0;
    for t=1:T
        product=1;
        for j=1:n
            product=product*(1-Pb(j))^(X(t,j));
        end
        summation=summation+Y(t)*X(t,i)*Pb(i)/(1-product);
    end
    P(i)=summation/sum(X(:,i));
end

------------------------main code--------------------------

%HW4.2 Ning Ma
```

```
X=importdata('X.dat.txt');
Y=importdata('Y.dat.txt');
n=size(X,2);
T=size(X,1);
P=zeros(n,1)+0.2;
FID=fopen('cse_hw4_2_likelihoodtable', 'w+');
L=cse_hw4_2_L(X,Y,P,n,T);
fprintf(FID,'%-2.1f %-5.1f \n',0,L);
for iteration=1:64
    P=cse_hw4_2_P(X,Y,P,n,T);
    L=cse_hw4_2_L(X,Y,P,n,T);
    fprintf(FID,'%-2.1f %-5.1f \n',iteration,L);
end
```