H.W. 6

Q.6).1)

a) $P(a,b \mid c,d) = P(a \mid c,d) \times P(b \mid a,c,d)$

$$= \frac{P(a,c,d)}{P(c)\,P(d \mid c)} \times \frac{P(a,b,c,d)}{P(a)\,P(c \mid a)\,P(d \mid a,c)}$$

$$= \frac{P(a)\,P(c \mid a)\,\cancel{P(d \mid a,c)}}{P(c)\,P(d \mid c)} \times \frac{P(a)\,P(b \mid a)\,P(c \mid a,b)\,P(d \mid c,b,c)}{\cancel{P(a)}\,\cancel{P(c \mid a)}\,\cancel{P(d \mid a,c)}}$$

$$= \frac{P(a)\,P(b \mid a)\,P(c \mid a,b)\,P(d \mid b,c)}{P(c)\,P(d \mid c)} \quad - (1)$$

Now, $P(d \mid c)$:-

$$= \sum_b P(b,d \mid c)$$

$$= \sum_b P(b \mid c)\,P(d \mid b,c) \quad - (2)$$

Further, $P(b \mid c)$:-

$$= \frac{P(c \mid b)\,P(b)}{\cancel{P(c)}} \quad - (3)$$

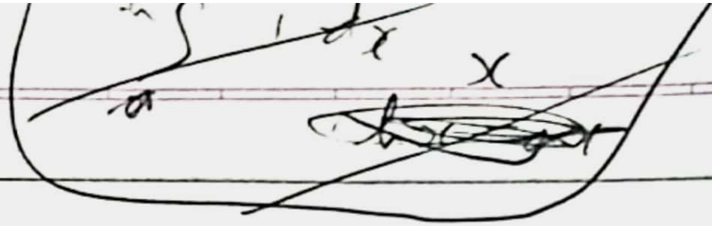
P(c) gets cancelled in (1) by P(c) in den.

To solve $P(c \mid b)$:-

$$= \sum_a P(a,c \mid b) = \sum_a P(a \mid b)\,P(c \mid a,b)$$

$$= \sum_a \frac{P(b \mid a)\,P(a)\quad P(c \mid a,b)}{\cancel{P(b)}} \quad - (4)$$

P(b) gets cancelled in (3) by P(b) in num.



Hence, using 1, 2, 3, 4 :-

$$P(a, b \mid c, d) = \frac{P(a)\, P(b \mid a)\, P(c \mid a, b)\, P(d \mid b, c)}{\sum_{b} \left(\sum_{a} \left(P(b \mid a)\, P(a)\, P(c \mid a, b) \right) P(d \mid b, c) \right)}$$

Q6). 1)

b)

i) $P(a \mid c,d)$

$$= \sum_b P(a, b \mid c, d) \quad \because \text{Marginalizn}$$

ii) $P(b \mid c,d)$

$$= \sum_a P(a, b \mid c, d) \quad \because \text{Marginalizn}$$

c)

$$\mathcal{L} = \sum_t \log P(C = c_t, D = d_t)$$

$$= \cancel{\sum_t \log P(C=c)\, P\, D=d}$$

$$= \sum_t \log P(C = c_t)\, P(D = d_t \mid C = c_t)$$

$$= \sum_t \log \sum_b \sum_a P(b \mid a)\, P(a)\, P(c_t \mid a, b)\, P(d_t \mid b, c_t)$$

- From (i) denominator.

d) i) $P(A = a)$

= For root node:-

$$= \frac{1}{T} \sum_{t=1}^{T} P(X_i = x \mid V_t = v_t)$$

$$= \frac{1}{T} \sum_{t=1}^{T} P(a \mid c_t, d_t) = \left(\frac{1}{T} \sum_{t=1}^{T}\right) //$$

Scanned with CamScanner

ii) $P(B=b \mid A=a)$

$$= \cancel{\sum_{t=1}^{T} P(b \mid a}$$

$$= \frac{\sum_t P(b, a \mid c_t, d_t)}{\sum_t P(a \mid c_t, d_t)}$$

$$= \frac{\sum_t P(a \mid c_t, d_t)\, P(b \mid c_t, d_t, a)}{\sum_t P(a \mid c_t, d_t)} \quad \because \text{Prod. rule.}$$

$$= \frac{\sum_t I(b, b_t)\, P(a \mid c_t, d_t)}{\sum_t P(a \mid c_t, d_t)}$$

//

iii) $P(C=c \mid A=a, B=b)$

$$= \frac{\sum_t P(c, a, b \mid c_t, d_t)}{\sum_t P(a, b \mid c_t, d_t)}$$

$$= \frac{\sum_t P(a, b \mid c_t, d_t)\, P(c \mid a, b, c_t, d_t)}{\sum_t P(a, b \mid c_t, d_t)}$$

$$= \frac{\sum_t I(c, c_t)\, P(a, b \mid c_t, d_t)}{\sum_t P(a, b \mid c_t, d_t)}$$

//

iv) $P(D=d \mid B=b, C=c)$

$$= \frac{\sum_t P(D=d, B=b, C=c \mid ct, dt)}{\sum_t P(B=b, C=c \mid ct, dt)}$$

$$= \frac{\sum_t P(d, b, c \mid ct, dt)}{\sum_t P(b, c \mid ct, dt)}$$

$$= \frac{\sum_t P(b \mid ct, dt)\, P(c \mid b, ct, dt)\, P(d \mid b, ct, c, dt)}{\sum_t P(b \mid ct, dt)\, P(c \mid b, ct, dt)}$$

$\because$ Prod rule

$$= \frac{\sum_t I(c, ct)\, I(d, dt)\, P(b \mid ct, dt)}{\sum_t I(c, ct)\, P(b \mid ct, dt)}$$

Q.6.2)

$p_i = P(z_i = 1 \mid x_i = 1)$

$0 = P(z_i = 1 \mid x_i = 0)$

& logical or for $P(y \mid z)$

a) $P(y=1 \mid x) = \sum_{z \in \{0,1\}^n} P(y=1, z \mid x)$

$= \sum_z P(z \mid x) \; P(y=1 \mid z, x)$

∵ Prod. rule.

$= \sum_z P(z \mid x) \; P(y=1 \mid z)$

∵ CI by d(i).

$= \sum_z P(z \mid x) \prod_{i=1}^{N} I(z_i, 1)$

∵ when $z=0$, $P(y=1 \mid z)$ is 0.

& 1 if $z=1$

$= \sum_z P(z_1, z_2, z_3 \dots \mid x_1, x_2, x_3 \dots) \prod_{i=1}^{N} I(z_i, 1)$

$= \prod_{i=1}^{N} P(z_i \mid x_i) \prod_{i=1}^{N} I(z_i, 1)$

∵ CI by d(iii)

$= \prod_{i=1}^{N} P(z_i = 1 \mid x_i)$

∵ when $z_i = 0$, whole term is 0.

$= 1 - \prod_{i=1}^{N} (1-p_i)^{x_i}$

$= P(y=1 \mid x)$

//

HP

h) $P(Z_i=1, X_i=1 \mid X=x, Y=y)$

$$= \frac{y\, x_i\, p_i}{1-\prod_j (1-p_j)^{x_j}}$$

LHS

$\therefore P(Z_i=1, X_i=1 \mid X=x, Y=y)$

$$= \frac{P(\cancel{X=x},\ Y=y \mid \overset{X=x,}{\wedge} Z_i=1, X_i=1)\ P(Z_i=1, X_i=1 \mid X=x)}{P(Y=y \mid X=x)}$$

$$= \frac{P(Y=y \mid Z_i=1)\ P(Z_i=1 \mid X=x)\ P(X_i=1 \mid X=x, Z_i=1)}{P(Y=y \mid X=x)}$$

$\because$ Cond I & Prod Rule

$$= \frac{I(y,1)\ I(x_i,1)\ P(Z_i=1 \mid X_i=x_i)}{P(Y=y \mid X=x)}$$

Now when y or $x = 0$, whole $=0$

else if y and $x_i=1$,

$$\frac{P(Z_i=1 \mid X_i=1)}{P(Y=1 \mid X=x)}$$

$\therefore \quad 0 \qquad y$ or $x_i=0$

$$\frac{p_i}{1-\prod_j (1-p_j)} \qquad y \,\&\, x_i=1$$

$$= \frac{y\, x_i\, p_i}{1-\prod_j (1-p_j)^{x_j}}$$

c)

$$p_i \leftarrow \frac{1}{T_i} \sum_t P(z_i=1, x_i=1 \mid X=x^{(t)}, Y=y^{(t)})$$

$\therefore P(z_i=1 \mid x_i=1)$

$$\leftarrow \frac{\sum_{t=1} P(z_i=1, x_i=1 \mid (X_t=x_t), (Y_t=y_t))}{\sum_{t=1} P(X_i=1 \mid X_t=x_t, Y_t=y_t)}$$

$$= \frac{\sum_{t=1} P(z_i=1, x_i=1 \mid X=x^{(t)}, Y=y^{(t)})}{\sum_t I(x_i^{(t)}, 1)}$$

$$= \sum_{t=1} P(z_i=10, x_i=1 \mid X=x^{(t)}, Y=y^{(t)}) \quad //$$

HP

Q.6) 3)

c

a) $f(x) = \log \cosh(x)$

$\therefore f'(x) = \frac{d}{dx} \log \cosh(x)$

$= \frac{1}{\cosh x} \times \frac{d}{dx} \cosh(x)$

$= \frac{\sinh(x)}{\cosh x} \times 1$

$\therefore f'(x) = \tanh(x)$

Now, if $f'(x) = 0$,

$\tanh(x) = 0$

$x = \tanh^{-1}(0)$

$\therefore x = 0$ — (1)

Now, $f''(x) = \frac{d}{dx} \tanh(x)$

$= \operatorname{sech}^2(x) \times 1$

Using (1),

$\therefore f''(x) = \operatorname{sech}^2(0)$

$= \frac{1}{\cosh^2(0)} = \frac{1}{1} = 1$

$\because 1 > 0 \therefore f''(x) > 0$

$\therefore$ Min occurs at $x = 0$

b) Now, from (a)

$$f''(x) = \operatorname{sech}^2(x)$$

$\because \operatorname{sech}(x) \in (0,1]$

when $x \in (-\infty, \infty)$

$\therefore \operatorname{sech}^2(x) \in (0,1]$

when $x \in (-\infty, \infty)$

$\therefore f''(x)$ is always $<= 1$

c) Plot In Pdf.

d)

i) $Q(x,x) = f(x)$

LHS :- $Q(x,x)$

$$= f(x) + f'(x)(x-x) + \frac{1}{2}(x-x)^2$$

$$= f(x) + 0 + 0$$

$$= f(x)$$

HP

Q:6)

i)

ii) now, $f(x) = f(y) + \int_y^x du \left[f'(y) + \int_y^u dv\, f''(v) \right]$

Now, from (h) $f''(v) \le 1$

$\therefore f(x) \le f(y) + \int_y^x du \left[f'(y) + \int_y^u 1\, dv \right]$

$\therefore f(x) \le f(y) + \int_y^x du \left[f'(y) + (u-y) \right]$

$\le f(y) + \int_y^x (f'(y) + u - y)\, du$

$\le f(y) + \left[u f'(y) + \frac{u^2}{2} - yu \right]_y^x$

$\le f(y) + xf'(y) + \frac{x^2}{2} - xy$

$- yf'(y) - \frac{y^2}{2} + y^2$

$\therefore f(x) \le f(y) + f'(y)(x-y) + \frac{1}{2}(x-y)^2$

$\therefore Q(x,y) \ge f(x)$ //

e) $x_{n+1} = \arg\min_x Q(x, x_n)$

$$= \arg\min_x \left(f(y) + f'(y)(x-y) + \frac{1}{2}(x-y)^2 \right)$$

$$= \arg\min_x \left(f(x_n) + f'(x_n)(x-x_n) + \frac{1}{2}(x-x_n)^2 \right)$$

$$= \arg\min_x \left(f(x_n) + f'(x_n)x - f'(x_n)x_n + \frac{x^2}{2} - x x_n + \frac{x_n^2}{2} \right)$$

$$= \arg\min_x \left(f(x_n) + x(f'(x_n) - x_n) - f'(x_n)x_n + \frac{x^2}{2} + \frac{x_n^2}{2} \right)$$

$$= \arg\min_x \underbrace{\left(\frac{x^2}{2} + x(f'(x_n) - x_n) + \left(\frac{x_n^2}{2} + f x_n - f' x_n x_n \right) \right)}_{g(x)}$$

To minimize $g(x)$,

$g'(x) = x + f'(x_n) - (x_n)$

If $g'x = 0$

$\therefore x = -(f'(x_n) - (x_n))$

$\therefore \arg\min_x (g(x))$

$= -(f'(x_n) - (x_n))$

$= x_n - f'(x_n)$ // $= x_n - \tanh(x_n)$ - from (a)

CSE250A_Hw6

November 11, 2022

[ ]:
```python
import numpy as np
from tqdm.notebook import tqdm
import time
from prettytable import PrettyTable
import math
import matplotlib.pyplot as plt
```

0.1 Q2d)

0.1.1 Load dataset

[ ]:
```python
X = np.loadtxt("X.txt")
```

[ ]:
```python
X.shape
```

[ ]: (267, 23)

[ ]:
```python
Y = np.loadtxt("Y.txt")
```

[ ]:
```python
Y.shape
```

[ ]: (267,)

[ ]:
```python
def logLikli(p,x):
  tmp = (1-p)**x
  Proby0Gx = np.prod(tmp)
  Proby1Gx = (1-Proby0Gx)

  return Proby0Gx,Proby1Gx
```

[ ]:
```python
def Estep(p,x,y):
  numi = p*y*x

  deno = (1-p)**x
  deno = np.prod(deno)
  deno = 1-deno
  return numi/deno
```

[ ]:
```
pi = np.array([0.05]*X.shape[1])
```

[ ]:
```
print(pi)
```

```
[0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05
 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05 0.05]
```

[ ]:
```
TotSamples = []
for i in range(len(X[0])):
  TotSamples.append(np.sum(X[:,i]))
print(TotSamples[:5])
```

```
[119.0, 66.0, 105.0, 76.0, 108.0]
```

[ ]:
```
pi = np.array([0.05]*X.shape[1])
N = 256
mistakes = []
loglihood = []
allMistakes = []
allLogli = []
toPrint = [0]
k = 0
for i in range(N):
  mist = 0
  logli = 0
  estep = 0

  for sample in range(X.shape[0]):
    #llhood and mistakes
    Proby0Gx,Proby1Gx = logLikli(pi,X[sample])
    y = Y[sample]
    if y == 1:
      llhood = Proby1Gx
      if Proby1Gx < 0.5:
        mist+=1
    else:
      llhood = Proby0Gx
      if Proby0Gx < 0.5:
        mist+=1
    logli+=np.log(llhood)

    # estep
    estep+=Estep(pi,X[sample],y)
  #   print(Proby0Gx,Proby1Gx,logli,estep)
  #   break
  # break

  pi = estep/TotSamples
```

```
  allMistakes.append(mist)
  allLogli.append(logli)

  if 2**k == i+1 or i==0:
    if i!=0:
      toPrint.append(2**k)
    mistakes.append(mist)
    loglihood.append(logli/X.shape[0])
    k+=1
```

[ ]:
```
print("Table:")
x = PrettyTable()
x.add_column("Iteration", toPrint)
x.add_column("# Mistakes", mistakes)
x.add_column("LogLikelihood", loglihood)
print(x)
```

```
Table:
+-----------+------------+----------------------+
| Iteration | # Mistakes |    LogLikelihood     |
+-----------+------------+----------------------+
|     0     |    175     | -0.9580854082157914  |
|     2     |     56     | -0.49591639407753635 |
|     4     |     40     | -0.3779406836061008  |
|     8     |     44     | -0.3500255657709249  |
|     16    |     40     | -0.33584054521650464 |
|     32    |     37     | -0.3230508504922716  |
|     64    |     37     | -0.3149545210186029  |
|    128    |     36     | -0.3111781626120766  |
|    256    |     36     | -0.31016351482108506 |
+-----------+------------+----------------------+
```

0.2 Q3c)

[ ]:
```
def f(x):
  return math.log(math.cosh(x))

def fd(x):
  return math.tanh(x)
```

[ ]:
```
def Q(x,y):
  return f(y) + (fd(y)*(x-y)) + (0.5*((x-y)**2))
```

[ ]:
```
# x= np.linspace(-1,1,10000)
# print(len(x))
p1 = []
```

```
p2 = []
p3 = []
for i in range(-100000,100001):
  i = i/10000
  p1.append(f(i))
  p2.append(Q(i,-2))
  p3.append(Q(i,3))
```

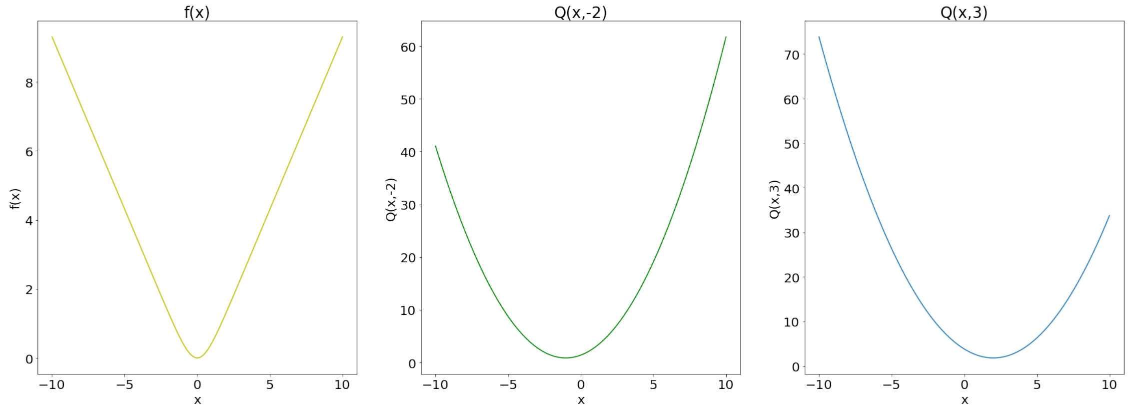
[ ]:
```
x= np.linspace(-10,10,200001)
# x=x/10000

len(x),len(p1)
```

[ ]: (200001, 200001)

[ ]:
```
# x= np.linspace(0/10000,10000/10000,10000)

# x = np.linspace(0,N,N)
fig=plt.figure(figsize=(30,10))
plt.rcParams.update({'font.size': 20})
fig.add_subplot(1,3,1)
plt.plot(x,p1,"y")
plt.xlabel('x')
plt.ylabel('f(x)')
plt.title('f(x)')
fig.add_subplot(1,3,2)
plt.plot(x,p2,"g")
plt.xlabel('x')
plt.ylabel('Q(x,-2)')
plt.title('Q(x,-2)')
fig.add_subplot(1,3,3)
plt.plot(x,p3)
plt.xlabel('x')
plt.ylabel('Q(x,3)')
plt.title('Q(x,3)')
plt.show()
```

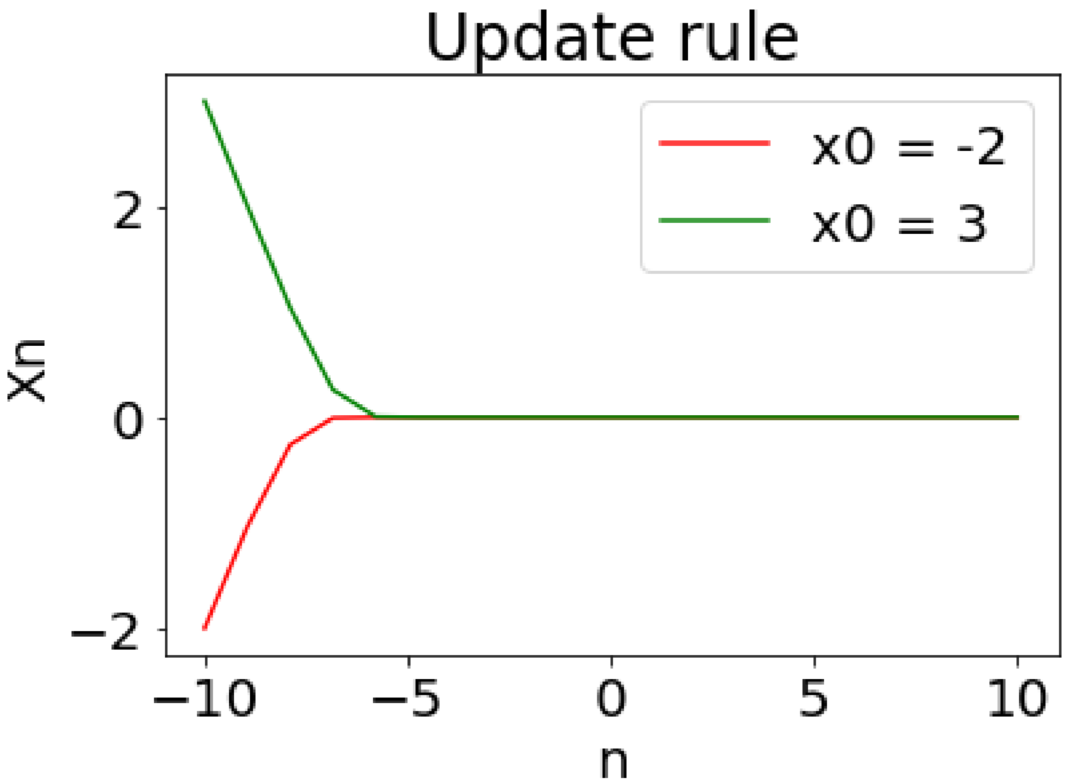


0.3 Q3d)

[ ]:
```
def updatee(xn):
  return xn-fd(xn)

def doUpdatee(x0,n):
  xn = [x0]
  for i in range(n-1):
    xn.append(updatee(xn[-1]))
  return xn
# x0 = -2 and x0 = 3
x = np.linspace(-10,10,20)
X0_1 = doUpdatee(-2,len(x))
X0_2 = doUpdatee(3,len(x))

plt.plot(x, X0_1, "r",label="x0 = -2")
plt.plot(x, X0_2, "g",label="x0 = 3")
plt.title('Update rule')
plt.ylabel('Xn')
plt.xlabel('n')
plt.legend()
plt.show()
```


Update rule
x0 = -2
x0 = 3
2
0
−2
Xn
−10
−5
0
5
10
n

```
[ ]:
```