5.1

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

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

$$= \frac{P(d|b,c)\,P(c|a,b)\,P(b|a)\,P(a)}{\sum_b P(d|b,c)\left(\sum_a P(c|a,b)\,P(a|b)\right)P(b)}$$

$$\boxed{P(a,b|c,d) = \frac{P(d|b,c)\,P(c|a,b)\,P(b|a)\,P(a)}{\left(\sum_b P(d|b,c)\left[\sum_a P(c|a,b)\,P(a|b)\right]\right)\sum_a P(b|a)\,P(a)}}$$

b) $$P(a|c,d) = \sum_b P(a,b|c,d)\,P(c,d) = \sum_b P(a,b|c,d)\,P(d|c)\,P(c)$$

$$= \sum_b \left[P(a,b|c,d)\left(\sum_b P(d|b,c)\,P(b|c)\right)\right.$$

$$= \sum_b \left[P(a,b|c,d)\left(\sum_b P(d|b,c)\,\frac{P(c|b)\,P(b)}{P(c)}\right)\right.$$

$$\boxed{\sum_b \left[P(a,b|c,d)\left(\sum_b P(d|b,c)\left(\sum_a \frac{P(c|b,a)\,P(a|b)\,P(b|a)\,P(a)}{\sum_{a,b} P(c|b,a)\,P(b|a)\,P(a)}\right)\right.\right.}$$

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

$$= \sum_t \log\left(\sum_{a,b} P(C=c_t, D=d_t|a,b)\,P(a,b)\right)$$

$$= \sum_t \log\left(\sum_{a,b} \frac{P(a,b|C=c_t, D=d_t)\,P(C=c_t, D=d_t)}{P(a,b)}\,P(a,b)\right)$$

$$\boxed{\mathcal{L} = \sum_t \log\left(\sum_{a,b} P(a,b|C=c_t, D=d_t)\,P(C=c_t, D=d_t)\right)}$$

5.2

a) $P(Y=1|X) = \sum_{z \in \{0,1\}^n} P(Y=1, Z|X) = \sum_{z \in \{0,1\}^n} P(Y=1|Z)P(Z|X)$

$= \sum_{z \in \{0,1\}^n} (1 - P(Y=0|Z))P(Z|X)$ *the only time $P(Y=0|Z) < 1$ is when $Z = \vec{0}_n$

$= (1 - P(Y=0|Z))\, P(Z = \vec{0}_n | X)$

$= 1 - P(Z_1 = 0|X_1) \cdot \ldots \cdot P(Z_n = 0|X_n)$

$= 1 - (1 - P(Z_1 = 1|X_1)) \cdot \ldots \cdot (1 - P(Z_n = 1|X))$

$= 1 - (1-p_1)^{X_1} \cdot \ldots \cdot (1-p_n)^{X_n}$

$$\boxed{P(Y=1|X) = 1 - \prod_{i=1}^{n} (1-p_i)^{X_i}}$$

b) $P(Z_i = 1, X_i = 1 | X = x, Y = y) = \dfrac{P(Y=y, X_i = 1 | Z_i = 1, X = x)\, P(Z_i = 1 | X = x)}{P(Y=y|X=x)}$

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

$$\boxed{\dfrac{y \cdot x_i p_i}{1 - \prod_j (1-p_j)^{x_j}}}$$

c) $P(Z_i = 1 | X_i = 1) = p_i \leftarrow \dfrac{\sum_t P(Z_i = 1, X_i = 1 | X = x_t, Y = y_t)}{\sum_t (X_i = 1 | X = x_t, Y = y_t)}$

$$\boxed{p_i \leftarrow \dfrac{\sum_t P(Z_i = 1, X_i = 1 | X = x_t, Y = y_t)}{T_i}}$$

```
  it       M          L
   0      95    -0.6549
   1      49    -0.4500
   2      42    -0.3940
   4      43    -0.3616
   8      44    -0.3466
  16      40    -0.3343
  32      37    -0.3225
  64      37    -0.3148
 128      36    -0.3112
 256      36    -0.3102
>>
```

```
% 5.2d
clear; clc; close all

load('X.mat')
load('Y.mat')

X = hw5X1;
Y = hw5Y;
T = size(X,1);
n = size(X,2);
num_iter = 256;

p = zeros(num_iter+2,n);
M = zeros(num_iter+1,1);
L = zeros(num_iter+1,1);
Py0x = zeros(T,1);
Py1x = zeros(T,1);

p(1,:) = 2/n*ones(1,n);
count = sum(X,1);
disp('  it       M           L')
for it = 1:num_iter+1

    L(it) = 0;
    for t = 1:T

        Py0x(t) = 1;
        for i = 1:n
            Py0x(t) = Py0x(t)*(1-p(it,i))^X(t,i);
        end
        Py1x(t) = 1 - Py0x(t);

        for i = 1:n
            p(it+1,i) = p(it+1,i) + (1/count(i))*Y(t)*X(t,i)*p(it,i)/Py1x(t);
        end

        if ((Py1x(t) >= 0.5 && Y(t) == 0) || (Py1x(t) <= 0.5 && Y(t) == 1))
            M(it) = M(it) + 1;
        end

        if Y(t) == 1
            L(it) = L(it) + (1/T)*log(Py1x(t));
        else
            L(it) = L(it) + (1/T)*log(Py0x(t));
        end
    end

    if ((floor(log2(it-1)) == log2(it-1)) || (it-1 == 0))
        disp(sprintf('%4d%8d%12.4f',it-1,M(it),L(it)))
    end
end
```

5.3

a) $f(x) = \log\cosh(x) = \log\left(\frac{e^x+e^{-x}}{2}\right)$

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

$$= \frac{e^x-e^{-x}}{e^x+e^{-x}}$$

when $f'(x)=0$,

$e^x = e^{-x}$

$\ln e^x = \ln e^{-x}$

$x = -x$

$x = 0$

and $f''(x) > 0 \ \forall x \in \mathbb{R}$,

so

$x=0$ is the global minimum

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

$$= \frac{e^{2x}+2+e^{-2x} - (e^{2x}-2+e^{-2x})}{(e^x+e^{-x})^2}$$

$$= \frac{4}{(e^x+e^{-x})^2}$$

b) $f''(x) = \frac{4}{(e^x+e^{-x})^2}$ is maximized when $g(x) = (e^x+e^{-x})^2$ is minimized

$g'(x) = 2(e^x+e^{-x})(e^x-e^{-x}) = 2(e^{2x}-e^{-2x}) = 0$

$2(e^{4x}-1) = 0$

$e^{4x} = 1$

$x = 0 \leftarrow$ global min

$$f''(x) \leq \frac{4}{(1+1)^2}$$

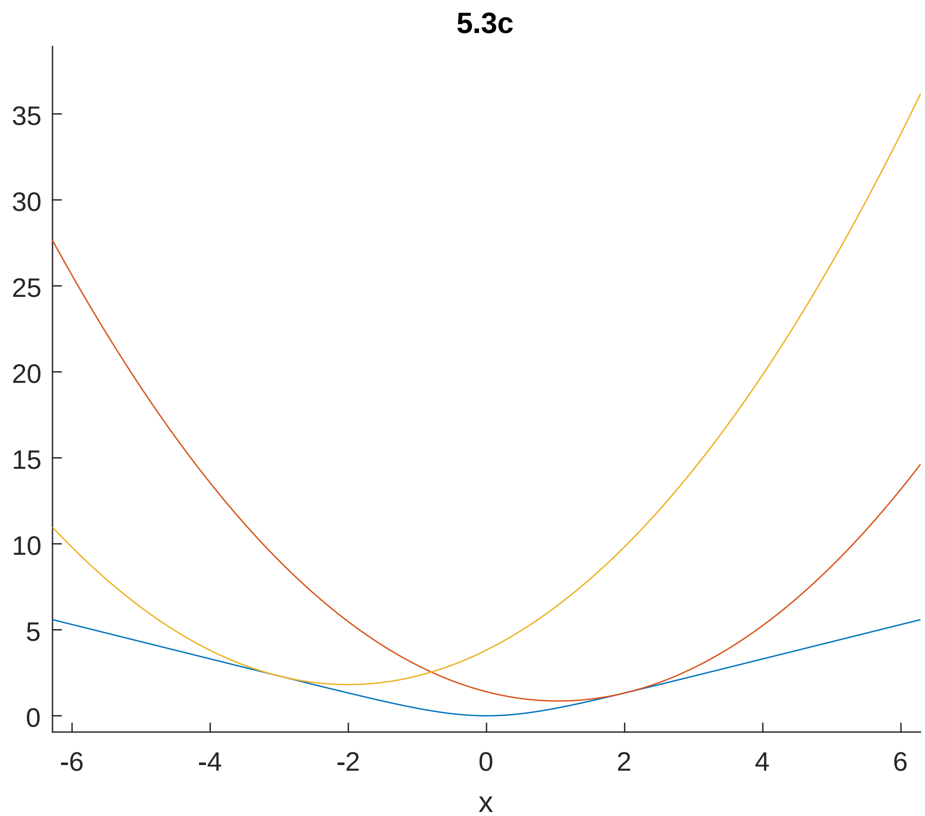
$$\boxed{f''(x) \leq 1}$$

c) see MATLAB plot

d) i) $Q(x,x) = f(x) + f'(x)\cancel{(x-x)} + \frac{1}{2}\cancel{(x-x)^2}$
$= f(x)$

ii) $Q(x,y) - f(x) = \log\left(\frac{e^y+e^{-y}}{2}\right) + \frac{e^y-e^{-y}}{e^y+e^{-y}}(x-y) + \frac{1}{2}(x-y)^2 - \log\left(\frac{e^x+e^{-x}}{2}\right)$

$$\boxed{Q(x,y) \geq f(x)}$$


5.3c
35
30
25
20
15
10
5
0
-6
-4
-2
0
2
4
6
x

```
% Prob 5.3c
clear; clc; close all

f = @(x) log(cosh(x));
df = @(x) (exp(x)-exp(-x))/(exp(x)+exp(-x));
Q1 = @(x) f(2) + df(2)*(x-2) + 0.5*(x-2)^2;
Q2 = @(x) f(-3) + df(-3)*(x+3) + 0.5*(x+3)^2;

hold on
ezplot(f)
% ezplot(df)
ezplot(Q1)
ezplot(Q2)
title('5.3c')
```

5.3

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

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

$\dfrac{\partial Q(x,x_n)}{\partial x} = f'(x_n) + (x - x_n) = 0$

$x = x_n - f'(x_n)$

$$\boxed{x_{n+1} = x_n - \frac{e^{x_n} - e^{-x_n}}{e^{x_n} + e^{-x_n}}}$$

f) see MATLAB code + plots

g) $x_{n+1} = x_n - \dfrac{f'(x_n)}{f''(x_n)} = x_n - \dfrac{\frac{e^{x_n} - e^{-x_n}}{e^{x_n} + e^{-x_n}}}{\frac{4}{(e^{x_n} + e^{-x_n})^2}} = x_n - \dfrac{(e^{x_n} - e^{-x_n})(e^{x_n} + e^{-x_n})}{4}$

$= x_n - \frac{1}{4}\left(e^{2x_n} - e^{-2x_n}\right)$

This converges when $|x_1| < |x_0|$

if $x_0 > 0$, then $-x_0 < x_1 < x_0$

$x_1 = x_0 - \frac{1}{4}\left(e^{2x_0} - e^{-2x_0}\right)$

$\Downarrow$

Newton's Method will diverge for large x_0

$e^{2x_0} - e^{-2x_0} > 0$

$e^{2x_0} > e^{-2x_0}$

$2x_0 > -2x_0$

✓

$e^{2x_0} - e^{-2x_0} < 8x_0$

$e^{2x_0} - e^{-2x_0} - 8x_0 < 0$

$\Downarrow$ MATLAB, $x_0 \in [0, 1.08]$

h) see MATLAB plot

The minimum occurs when $g'(x) = 0$

$g'(x) = \dfrac{1}{10}\sum_{k=1}^{10} \dfrac{1}{\cosh\left(x + \frac{1}{\sqrt{k^2+1}}\right)} \cdot \dfrac{e^{x + \frac{1}{\sqrt{k^2+1}}} - e^{-\left(x + \frac{1}{\sqrt{k^2+1}}\right)}}{2}$

$\Rightarrow$ Finding the minimum analytically is not simple

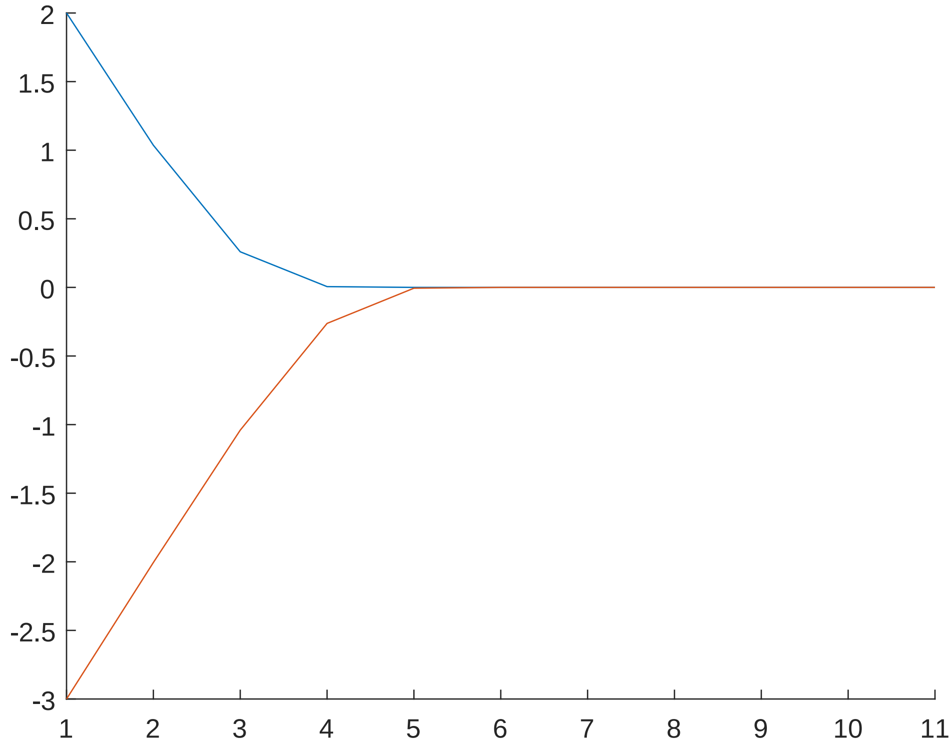
i) $R(x,x) = g(x) + g'(x)(x - x) + \frac{1}{2}(x - x)^2 = g(x)$ ✓

$R(x,y) = g(y) + g'(y)(x - y) + \frac{1}{2}(x - y)^2$

$=$

j) see MATLAB

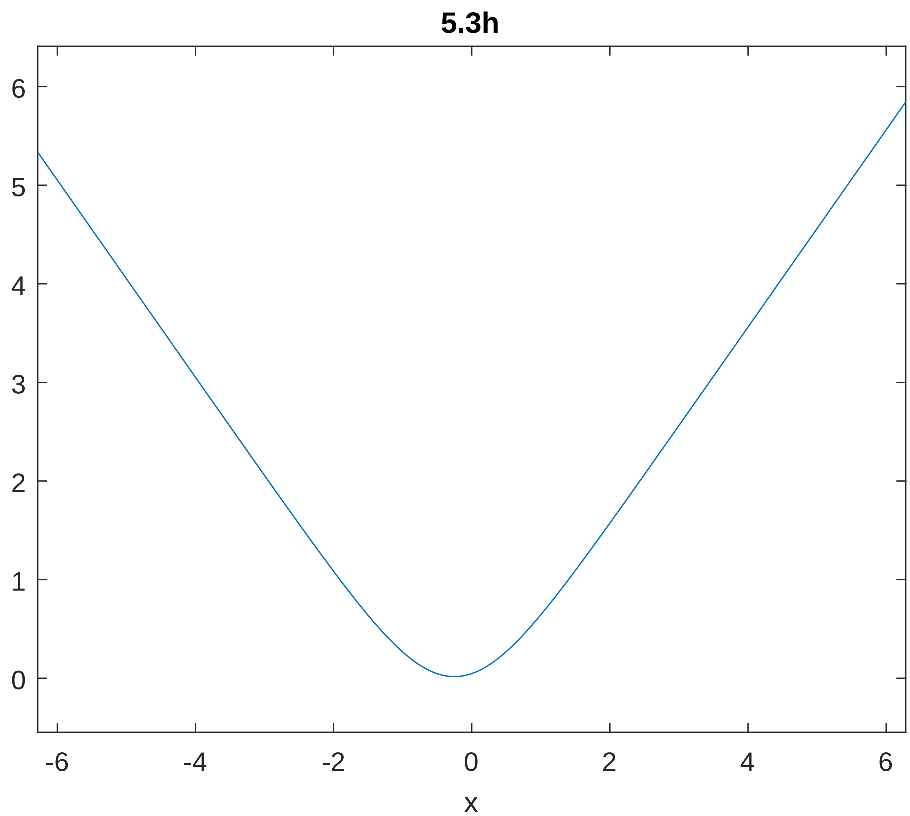
5.3f



```
% Prob % Prob 5.3f
clear; clc; close all

num_iter = 10;
x1 = zeros(1,num_iter);
x2 = zeros(1,num_iter);
x1(1) = 2;
x2(1) = -3;
for i = 1:num_iter
    x1(i+1) = x1(i) - (exp(x1(i))-exp(-x1(i)))/(exp(x1(i))+exp(-x1(i)));
    x2(i+1) = x2(i) - (exp(x2(i))-exp(-x2(i)))/(exp(x2(i))+exp(-x2(i)));
end

hold on
plot(x1)
plot(x2)
title('5.3f')
```


5.3h
6
5
4
3
2
1
0
-6
-4
-2
0
2
4
6
x

```
% Prob 5.3h
clear; clc; close all

g = @(x) 0;
for k = 1:10
    g = @(x) g(x) + 0.1*log(cosh(x+1/sqrt(k^2+1)));
end

ezplot(g)
title('5.3h')
```

5.3

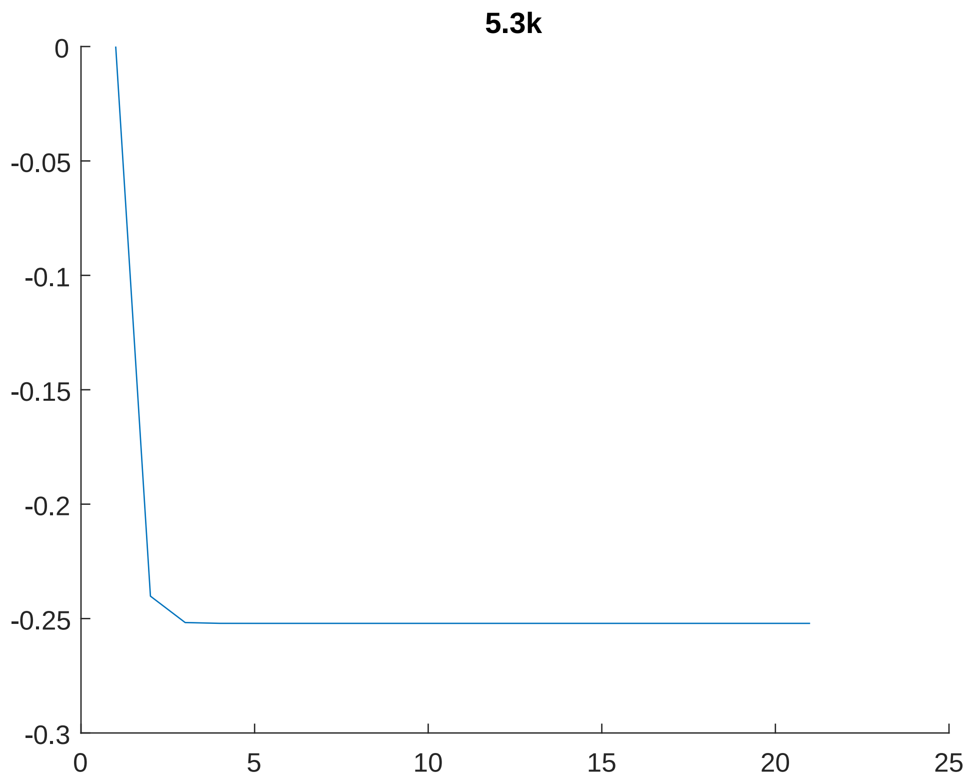
j) $x_{n+1} = \arg\min_{x} R(x, x_n)$ which is when $\frac{\partial R(x, x_n)}{\partial x} = 0$

$g'(x_n) + x - x_n = 0$

$\boxed{x_{n+1} = x_n - g'(x_n) \text{ , where } g'(x_n) \text{ is given in part h.}}$

k) See MATLAB code

$$\boxed{\begin{array}{l} x_{min} = -0.2521 \\ g(x) = 0.01624 \end{array}}$$


5.3k
0
-0.05
-0.1
-0.15
-0.2
-0.25
-0.3
0
5
10
15
20
25

```
% Prob % Prob 5.3k
clear; clc; close all

num_iter = 20;
x = zeros(1,num_iter);
x(1) = 0;
for i = 1:num_iter
    term = 0;
    for k = 1:10
        term = term + 0.05*(exp(x(i)+1/sqrt(k^2+1))-exp(-x(i)-1/sqrt(k^2+1)))./...
          cosh(x(i)+1/sqrt(k^2+1));
    end
    x(i+1) = x(i) - term;
end

hold on
plot(x)
disp(x)
title('5.3k')
```