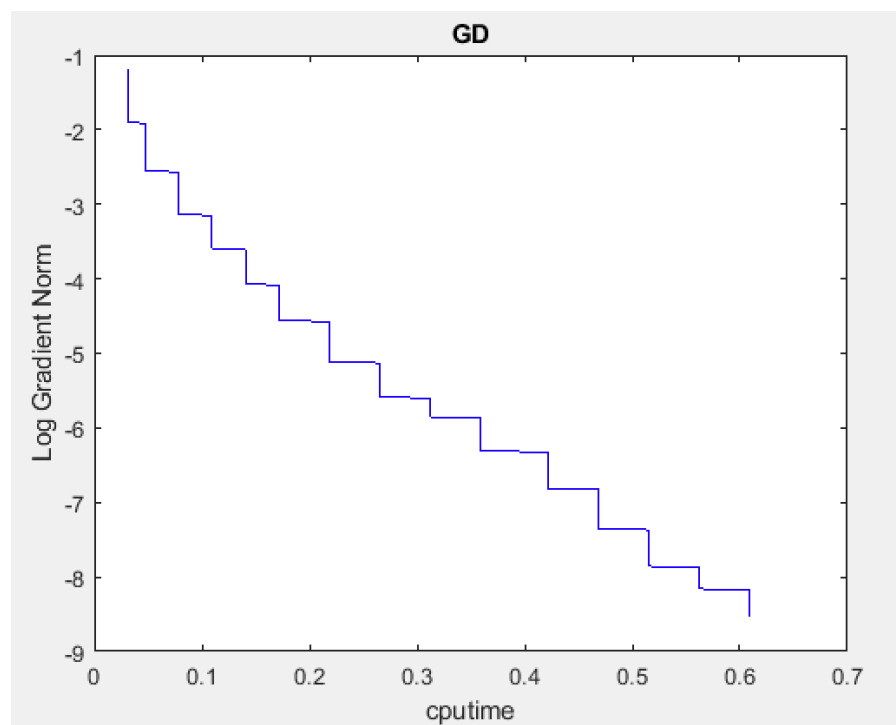
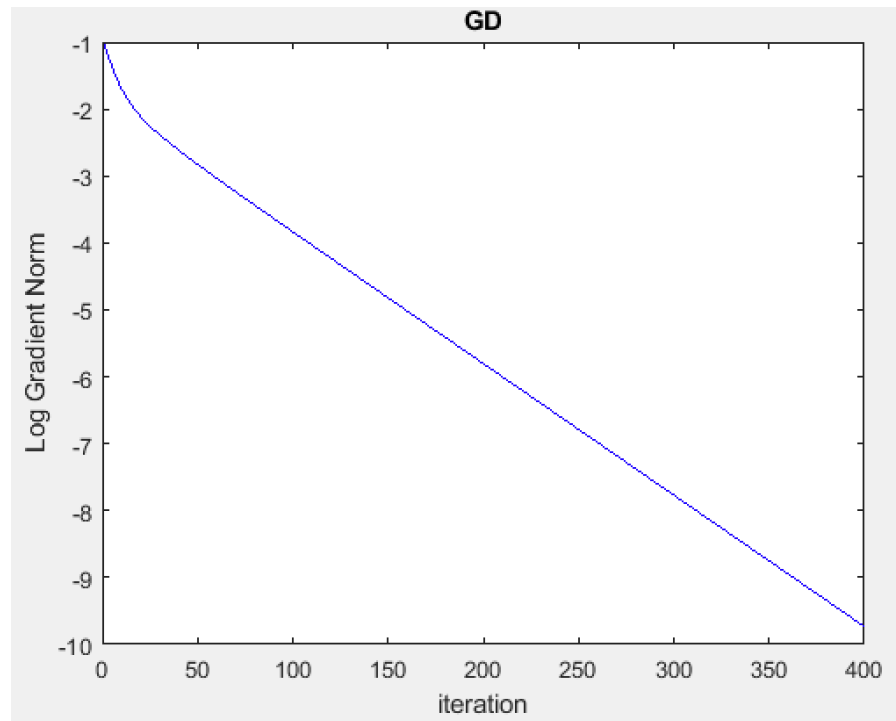
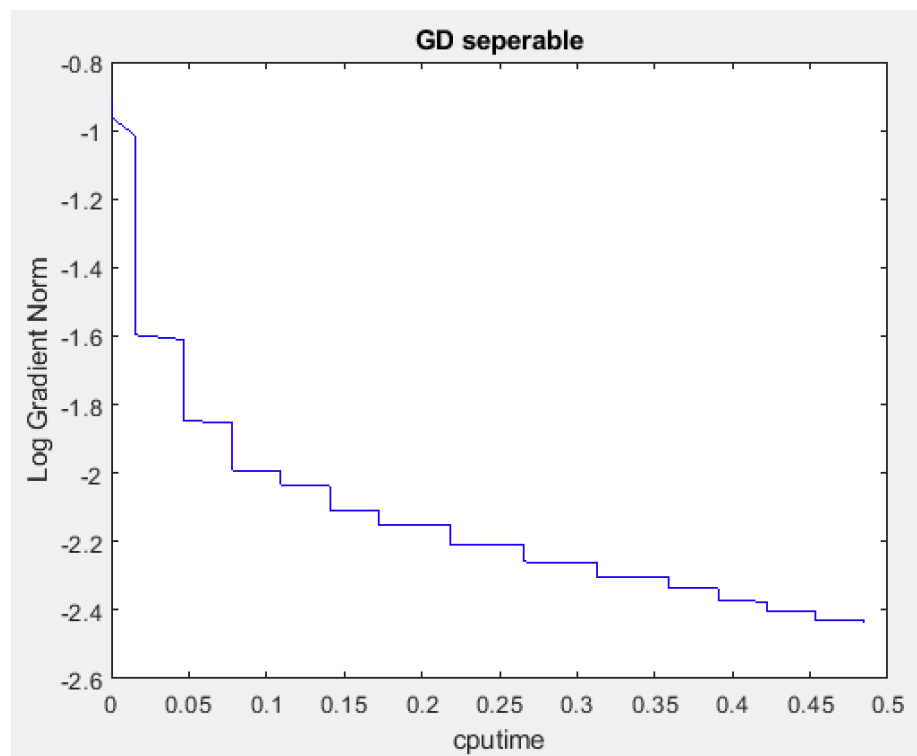
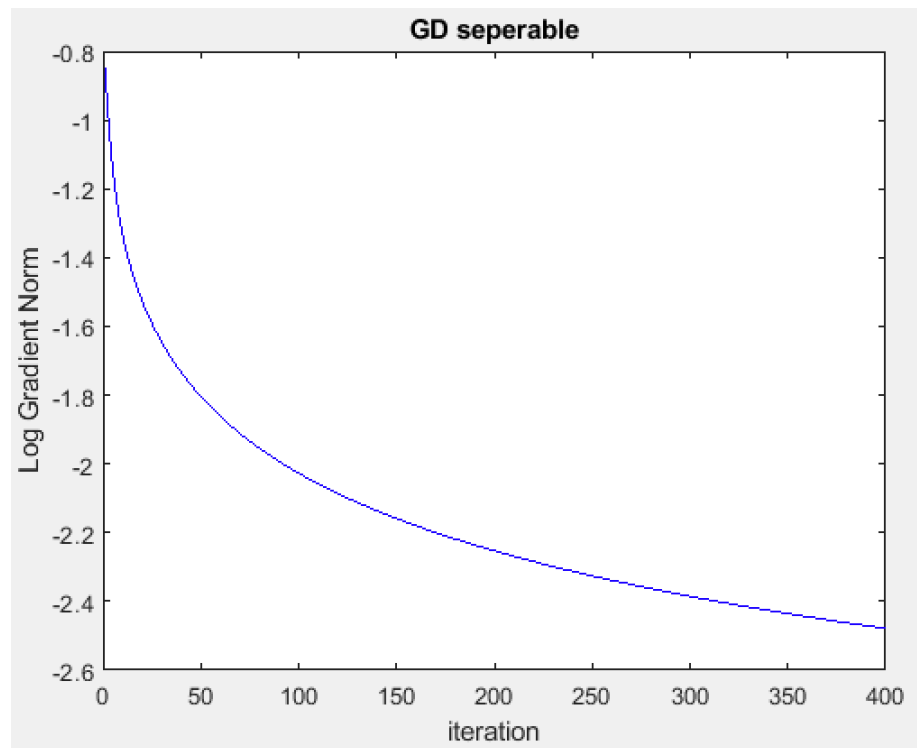


IE510 HW5
Problem 4

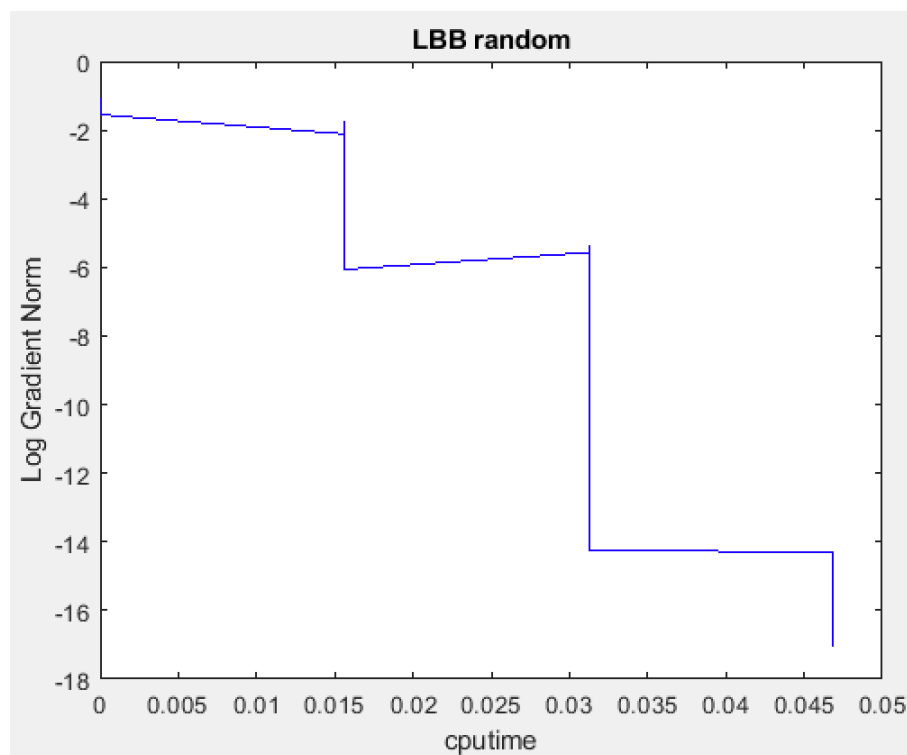
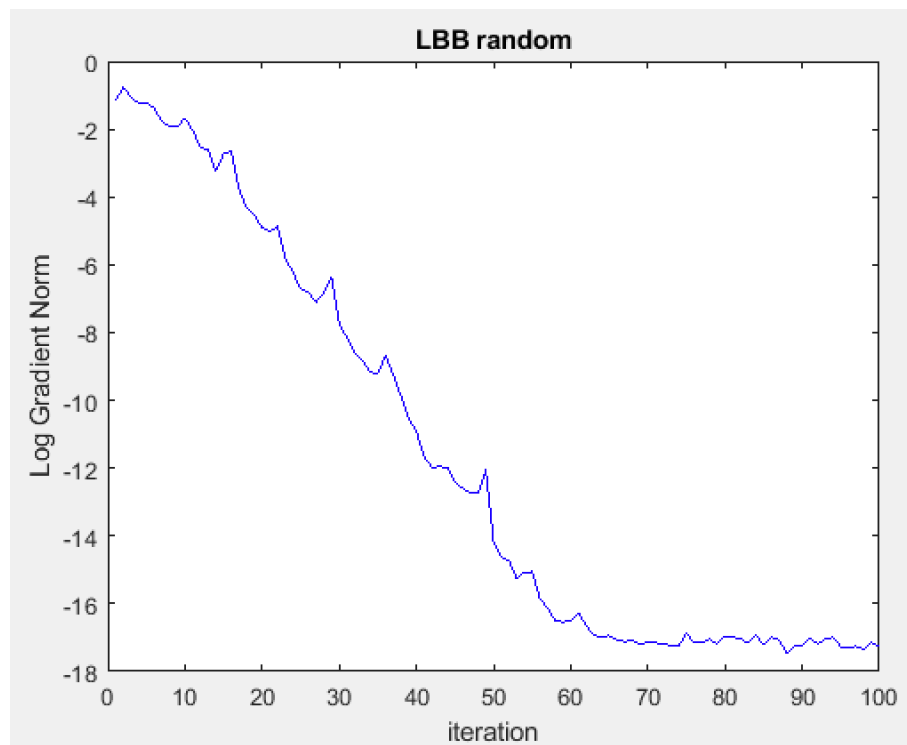
GD: random case



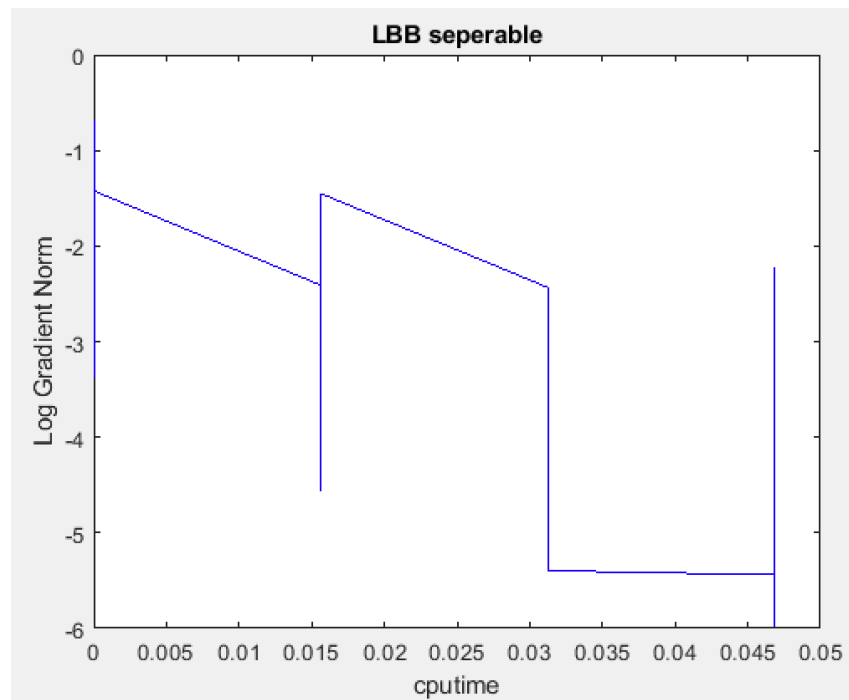
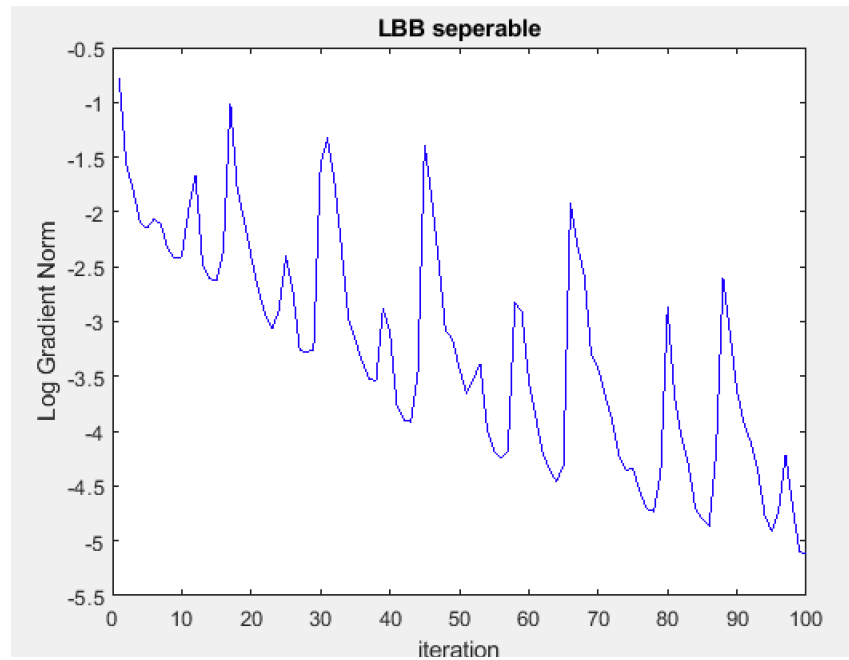
GD: separable case



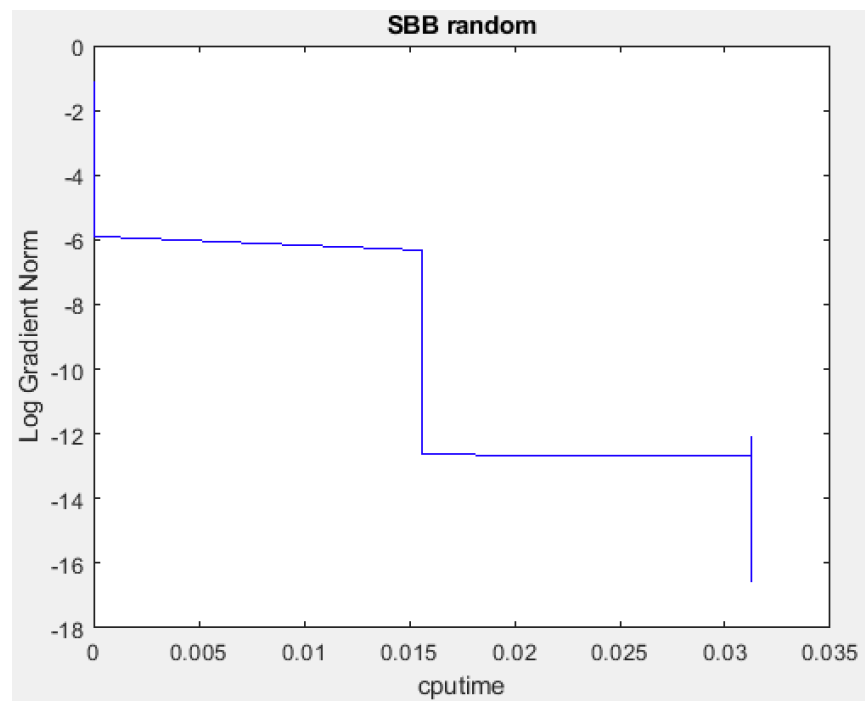
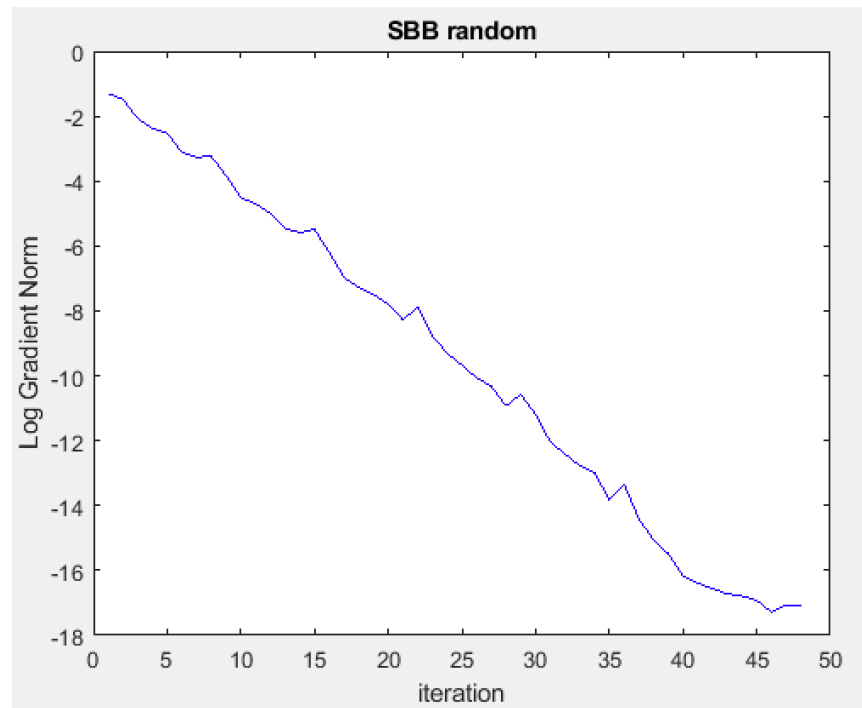
LBB: random case



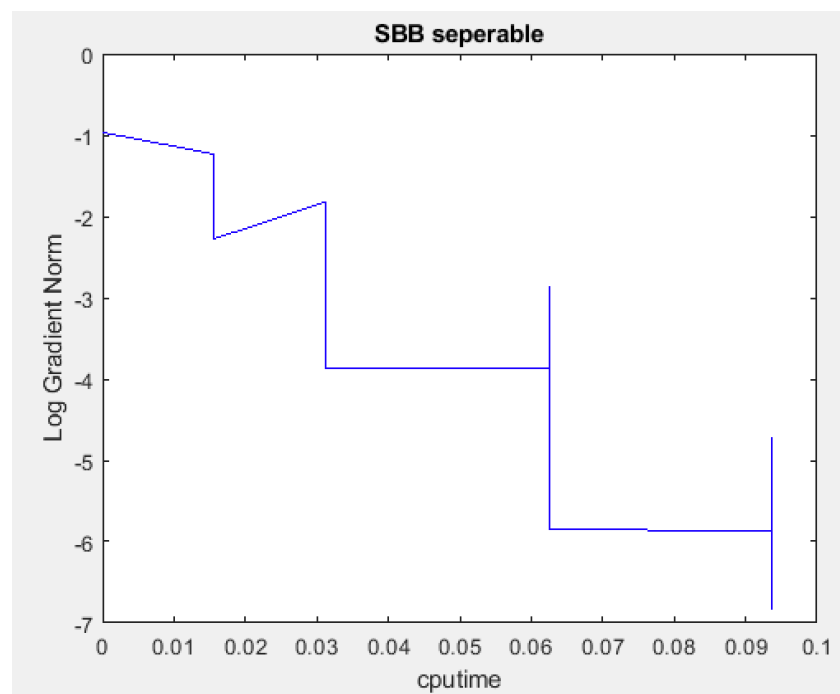
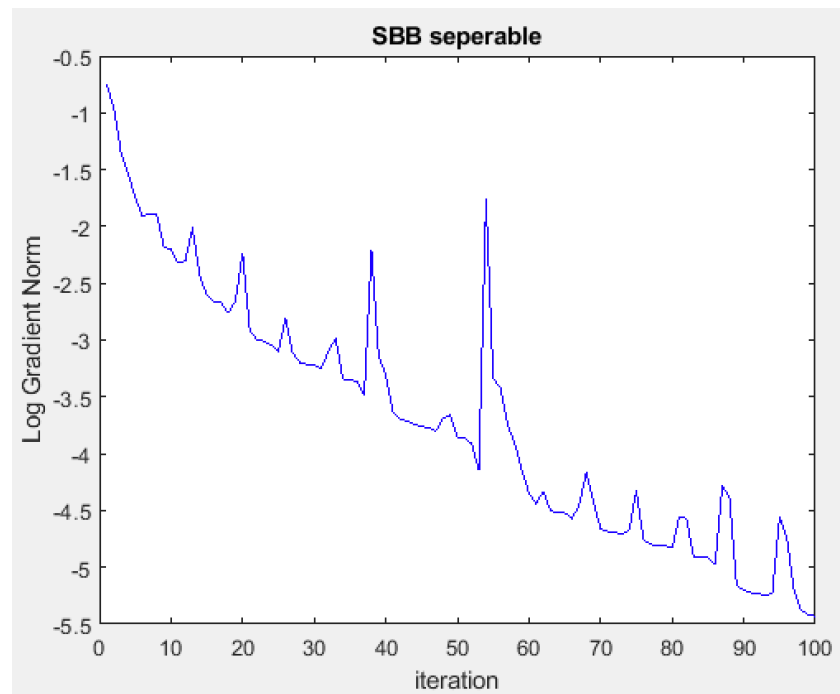
LBB: separable case



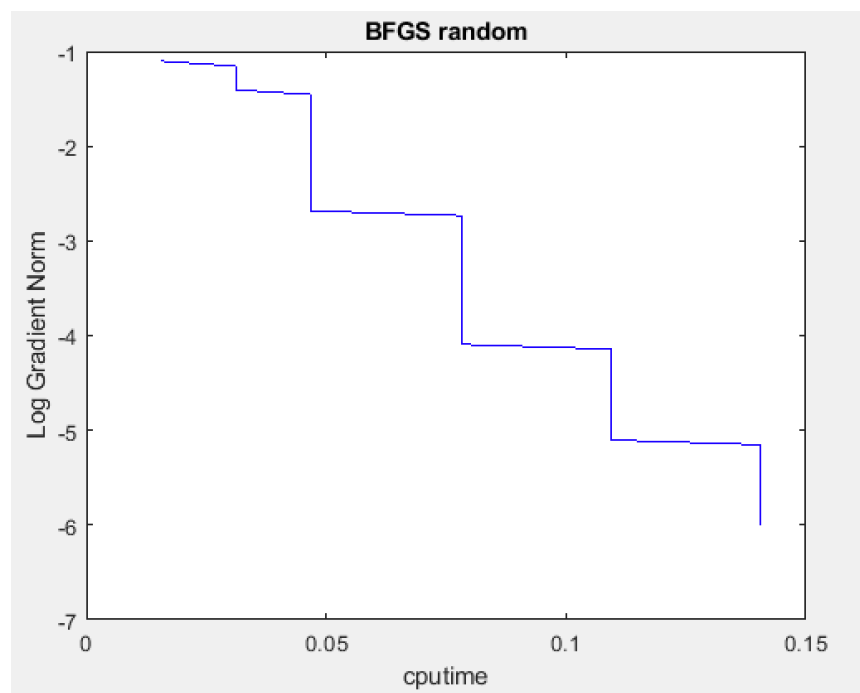
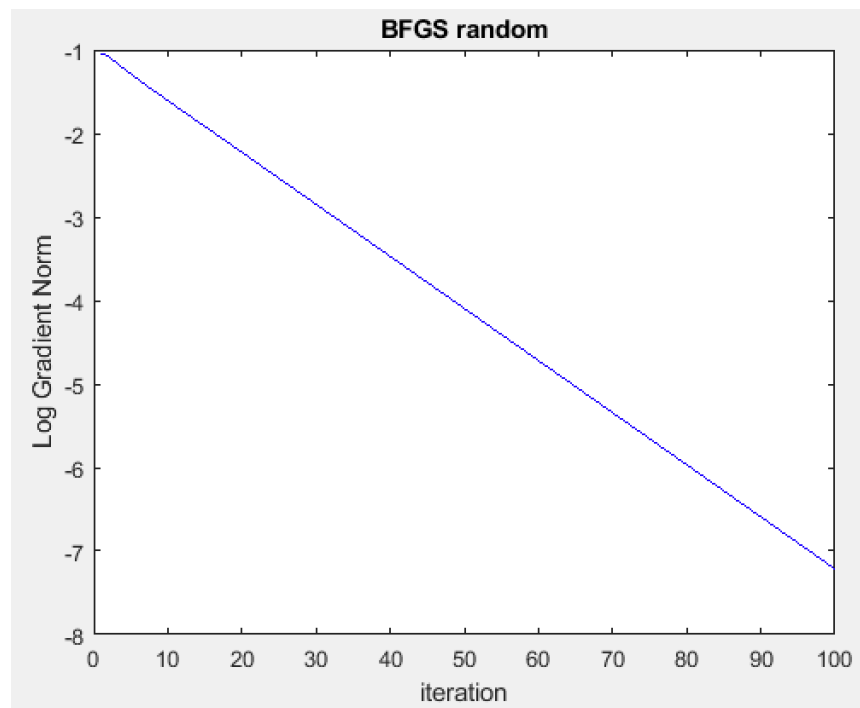
SBB: random case



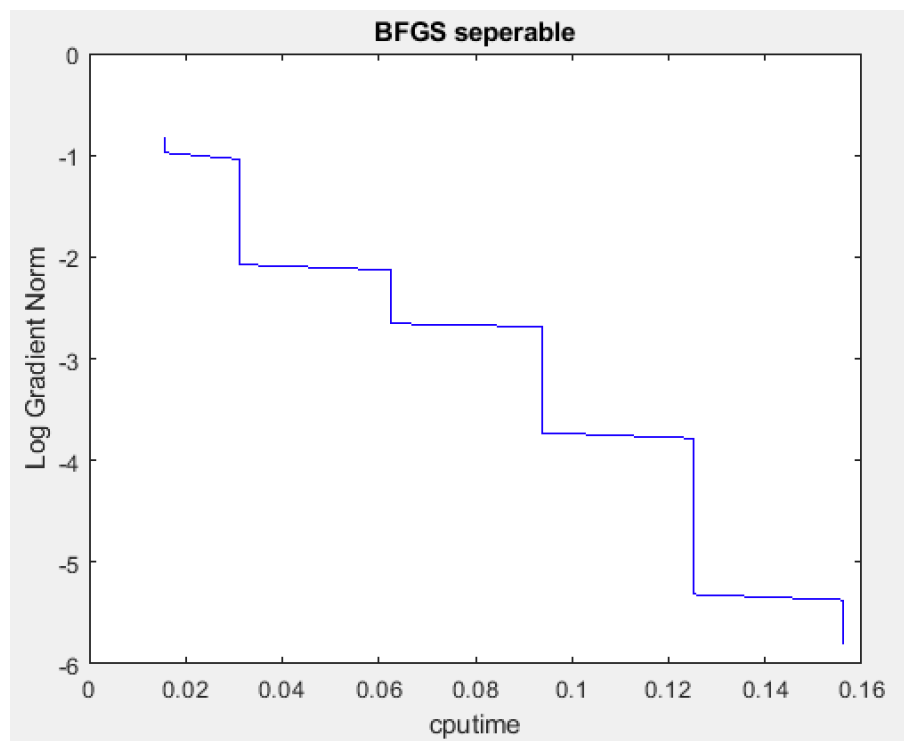
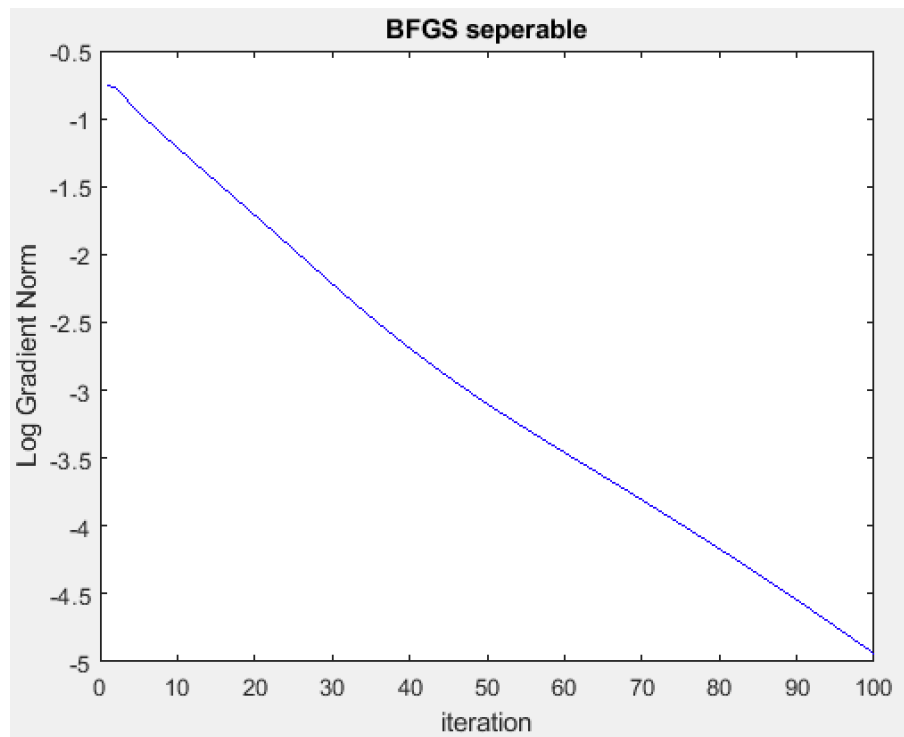
SBB: separable case



BFGS: random case



BFGS: separable case



Code: GD

```
clc;clear;
%%
n = 50;
d = 10;
X = rand(n,d)-0.5;
Y1 = binornd(1,0.5,n,1);
Y1(Y1==0)=-1;
w_star = ones(d,1);
Y2 = X*w_star;
Y2(Y2<0)=-1;
Y2(Y2>=0)=1;
Y = Y2;
%%
w = normrnd(0,1,[d,1]);
L = max(eig(X'*X));
max_it = 4*10^2;
grad_norm = zeros(max_it + 1, 1);

ite = 1;
alpha = 1/L;
t = cputime;
times = zeros(max_it + 1, 1);
while(ite<=max_it)
    grad_w = zeros(d,1);
    for i=1:n
        z = exp(-Y(i)*X(i,:)*w);
        grad_w = grad_w+ (-Y(i)*X(i,:)*z)'/ (1+z);
    end
    w = w - alpha*grad_w;
    grad_norm(ite) = log10(norm(grad_w/n));
    times(ite) = cputime-t;
    ite = ite +1;
end
figure;
plot_length = max_it;
plot_vec = 1:1:plot_length;
plot(plot_vec, grad_norm(1:plot_length), 'b-');
plot(times(1:plot_length), grad_norm(1:plot_length), 'b-');

xlabel('iteration');
xlabel('cputime');
ylabel('Log Gradient Norm');
title('GD seperable');
```

Code: LBB

```
%%clc;clear;
n = 50;
d = 10;
X = rand(n,d)-0.5;
Y1 = binornd(1,0.5,n,1);
Y1(Y1==0)=-1;
w_star = ones(d,1);
Y2 = X*w_star;
Y2(Y2<0)=-1;
Y2(Y2>=0)=1;
Y = Y2;
I = eye(d);
max_it = 10^2;
grad_norm = zeros(max_it + 1, 1);
w = normrnd(0,1,[d,1]);
w_prev = normrnd(0,1,[d,1]);
grad_prev = normrnd(0,1,[d,1]);
t = cputime;
times = zeros(max_it + 1, 1);
ite = 1;

while(ite<=max_it)
    grad_w = zeros(d,1);
    sum = 0;
    for i=1:n
        z = exp(-Y(i)*X(i,:)*w);
        grad_w = grad_w+ (-Y(i)*X(i,:)*z)'/ (1+z);
    end
    grad_norm(ite) = log10(norm(grad_w/n));
    s = w - w_prev;
    y = grad_w - grad_prev;
    alpha = (s'*s)/(y'*s);
    H_k = alpha*I;
    grad_prev = grad_w;
    w_prev = w;
    w = w - H_k*grad_w;
    times(ite) = cputime-t;
    ite = ite +1;
end

figure;
plot_length = max_it;
plot_vec = 1:1:plot_length;
plot(plot_vec, grad_norm(1:plot_length), 'b-');
%plot(times(1:plot_length), grad_norm(1:plot_length), 'b-');
xlabel('iteration');
%xlabel('cputime');
ylabel('Log Gradient Norm');
title('LBB seperable');
```

Code: SBB

```
%%clc;clear;
n = 50;
d = 10;
X = rand(n,d)-0.5;
Y1 = binornd(1,0.5,n,1);
Y1(Y1==0)=-1;
w_star = ones(d,1);
Y2 = X*w_star;
Y2(Y2<0)=-1;
Y2(Y2>=0)=1;
Y = Y1;
I = eye(d);
max_it = 10^2;
grad_norm = zeros(max_it + 1, 1);
w = normrnd(0,1,[d,1]);
w_prev = normrnd(0,1,[d,1]);
grad_prev = normrnd(0,1,[d,1]);
t = cputime;
times = zeros(max_it + 1, 1);
ite = 1;

while(ite<=max_it)
    grad_w = zeros(d,1);
    sum = 0;
    for i=1:n
        z = exp(-Y(i)*X(i,:)*w);
        grad_w = grad_w+ (-Y(i)*X(i,:)*z)/(1+z);
    end
    grad_norm(ite) = log10(norm(grad_w/n));
    s = w - w_prev;
    y = grad_w - grad_prev;
    alpha = (y'*s)/(y'*y);
    H_k = alpha*I;
    grad_prev = grad_w;
    w_prev = w;
    w = w - H_k*grad_w;
    times(ite) = cputime-t;
    ite = ite +1;
end

figure;
plot_length = max_it;
plot_vec = 1:1:plot_length;
plot(plot_vec, grad_norm(1:plot_length), 'b-');
plot(times(1:plot_length), grad_norm(1:plot_length), 'b-');
xlabel('iteration');
xlabel('cputime');
ylabel('Log Gradient Norm');
title('SBB random');
```

Code: BFGS

```
%%clc;clear;
n = 50;
d = 10;
X = rand(n,d)-0.5;
Y1 = binornd(1,0.5,n,1);
Y1(Y1==0)=-1;
w_star = ones(d,1);
Y2 = X*w_star;
Y2(Y2<0)=-1;
Y2(Y2>=0)=1;
Y = Y1;
I = eye(d);
max_it = 100;
L = max(eig(X'*X));
alpha = 1/L;
grad_norm = zeros(max_it + 1, 1);
w = normrnd(0,1,[d,1]);
w_prev = zeros(d,1);
grad_prev = zeros(d,1);
H_k = eye(d);
t = cputime;
times = zeros(max_it + 1, 1);
ite = 1;
while(ite<=max_it)
    grad_w = zeros(d,1);
    for i=1:n
        z = exp(-Y(i)*X(i,:)*w);
        grad_w = grad_w+ (-Y(i)*X(i,:)*z)'/ (1+z);
    end
    grad_norm(ite) = log10(norm(grad_w/n));
    s = w - w_prev;
    y = grad_w - grad_prev;
    p = 1/(s'*y);
    H_k = (I-p*s*y')*H_k*(I-p*y*s')+p*(s*s');
    grad_prev = grad_w;
    w_prev = w;
    w = w - alpha*H_k*grad_w;
    times(ite) = cputime-t;
    ite = ite +1;
end

figure;
plot_length = max_it;
plot_vec = 1:1:plot_length;
plot(plot_vec, grad_norm(1:plot_length), 'b-');
%plot(times(1:plot_length), grad_norm(1:plot_length), 'b-');
xlabel('iteration');
%xlabel('cputime');
ylabel('Log Gradient Norm');
title('BFGS seperable');
```