HW 6

Note: I only verify my source code in p3.py and gd.py.

1. a) $\nabla^2 f(x) = \frac{1}{2} \cdot 2 \cdot Q = Q$.

   By second-order condition, $-LI \leq Q \leq LI$.
   $-L \leq \gamma \leq L$. So, the smallest L is $|\gamma| = \gamma$.

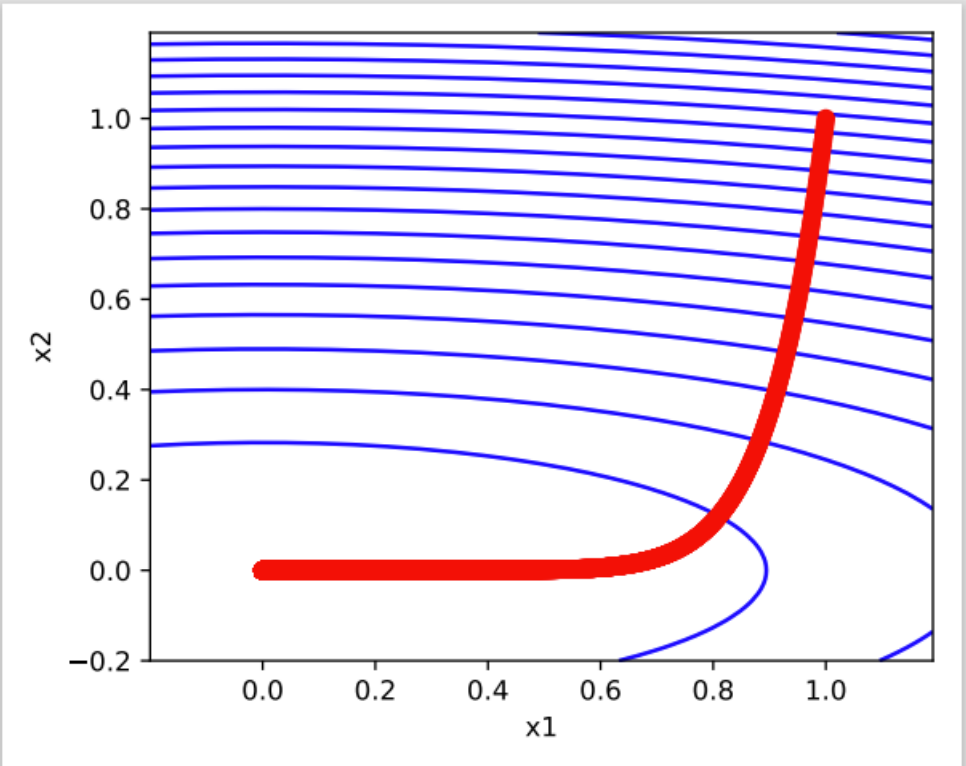
   b) $\tilde{f}(x) = \frac{1}{2} x^T Q x - \frac{m}{2} \|x\|^2 = \frac{1}{2} x^T Q x - \frac{m}{2} x^T x = \frac{1}{2} x^T (Q - mI) x$.

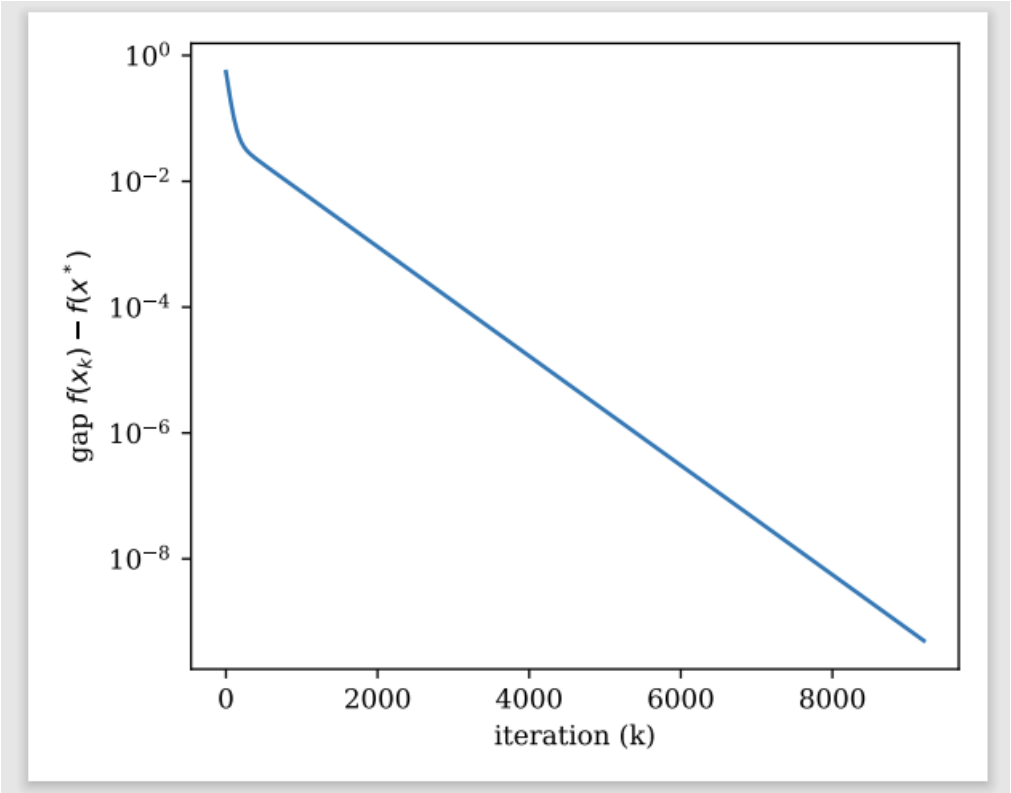
   Because $\tilde{f}$ is convex, $\nabla^2 \tilde{f}(x) = (Q - mI)$ is positive definite.
   So, the maxima of m is $\max\{1, \gamma\}$.

   c)

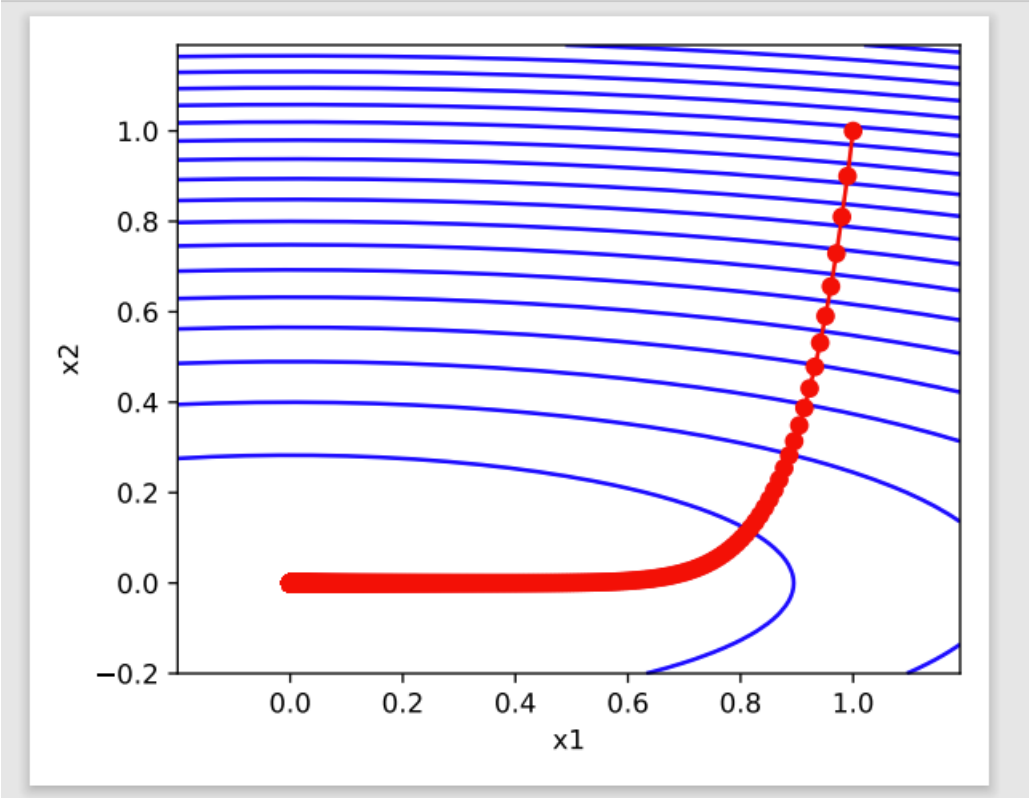
```
gamma=0.1, stepsize=2.2, number of iterations=5
gamma=0.1, stepsize=1, number of iterations=88
gamma=0.1, stepsize=0.1, number of iterations=917
gamma=0.1, stepsize=0.01, number of iterations=9206
```

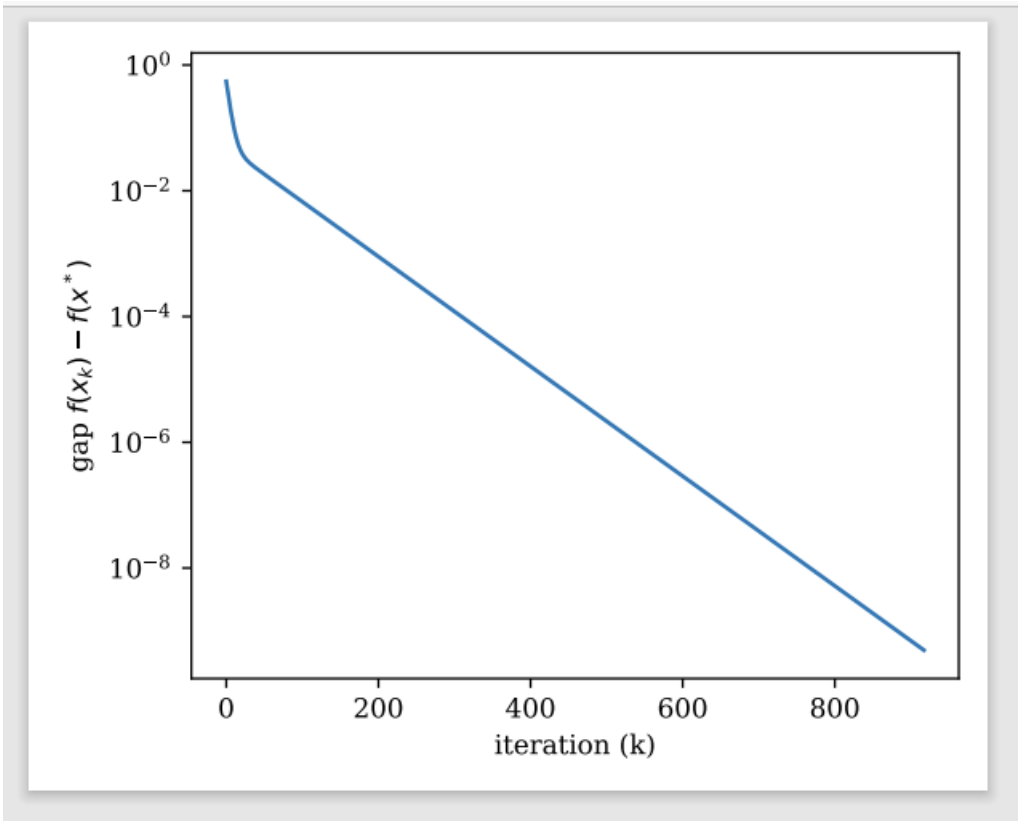
When taking 2.2 as the step size, it fails to converge.
ss = 0.01



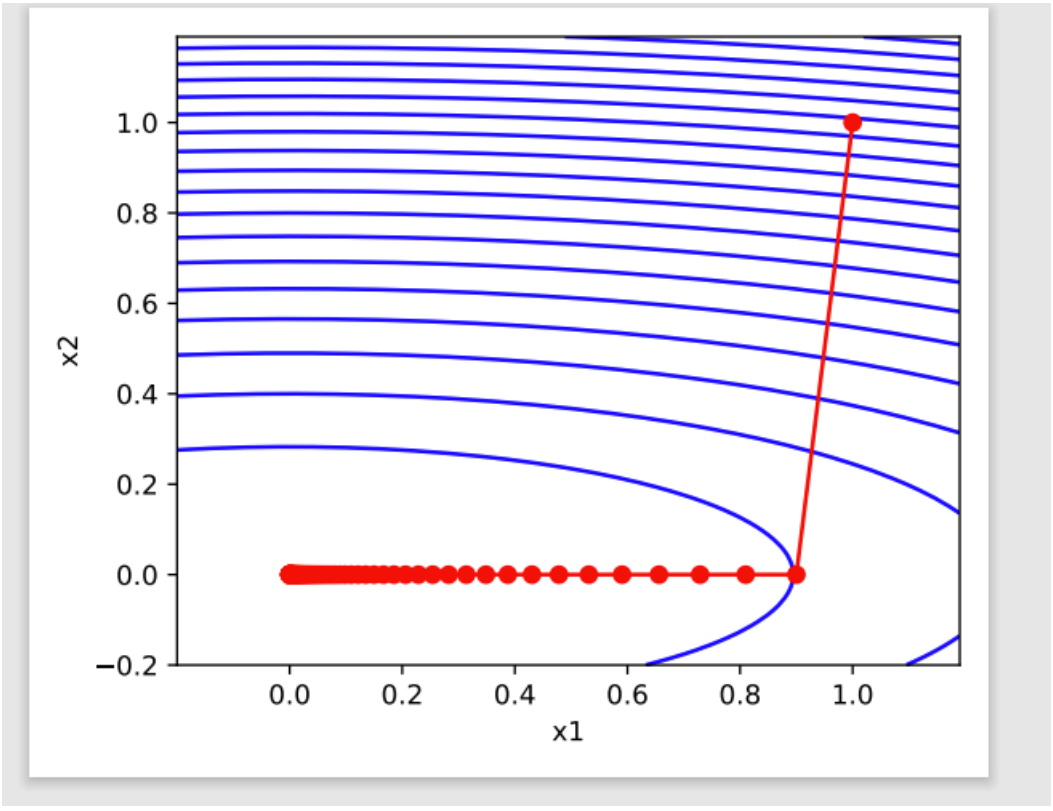

10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
gap $f(x_k) - f(x^*)$
0
2000
4000
6000
8000
iteration (k)

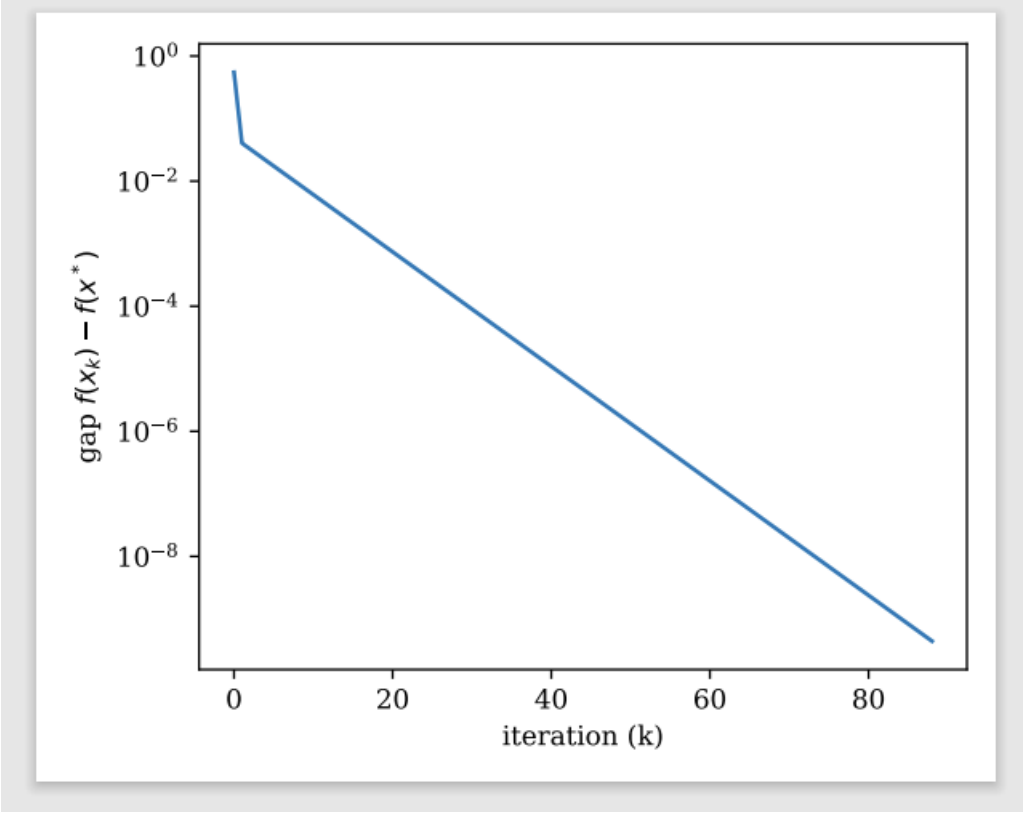
Ss = 0.1


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1

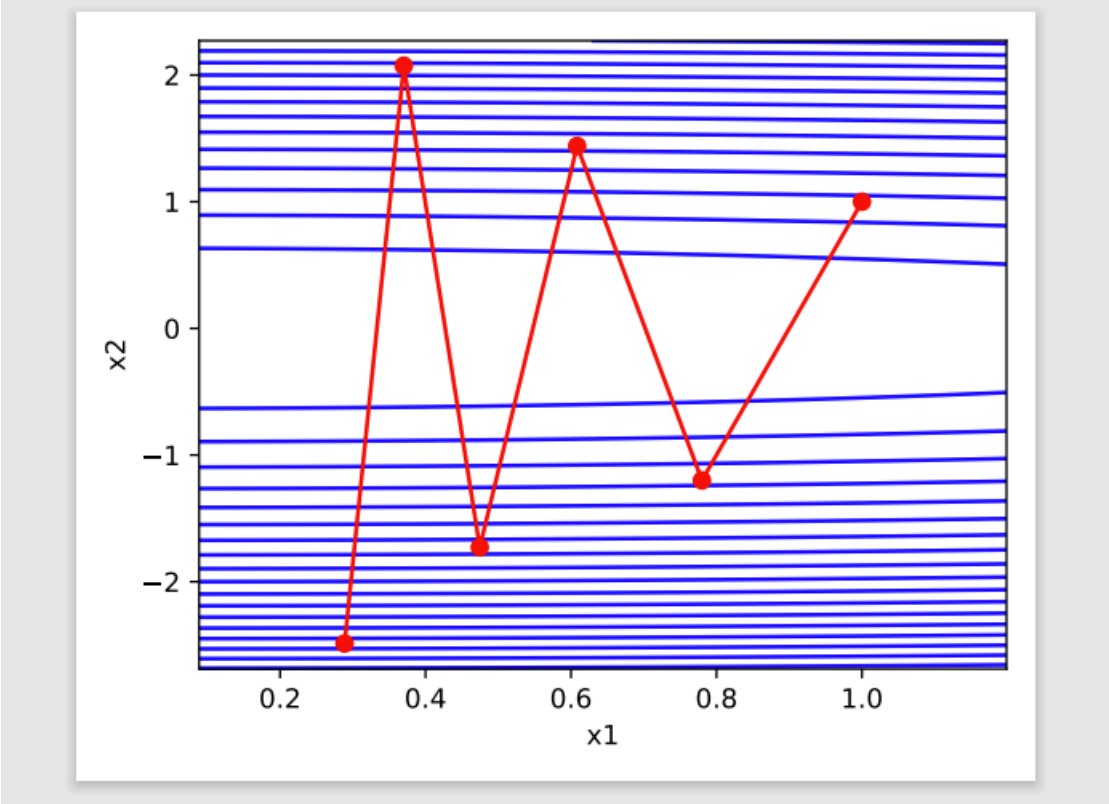

10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
gap $f(x_k) - f(x^*)$
0
200
400
600
800
iteration (k)

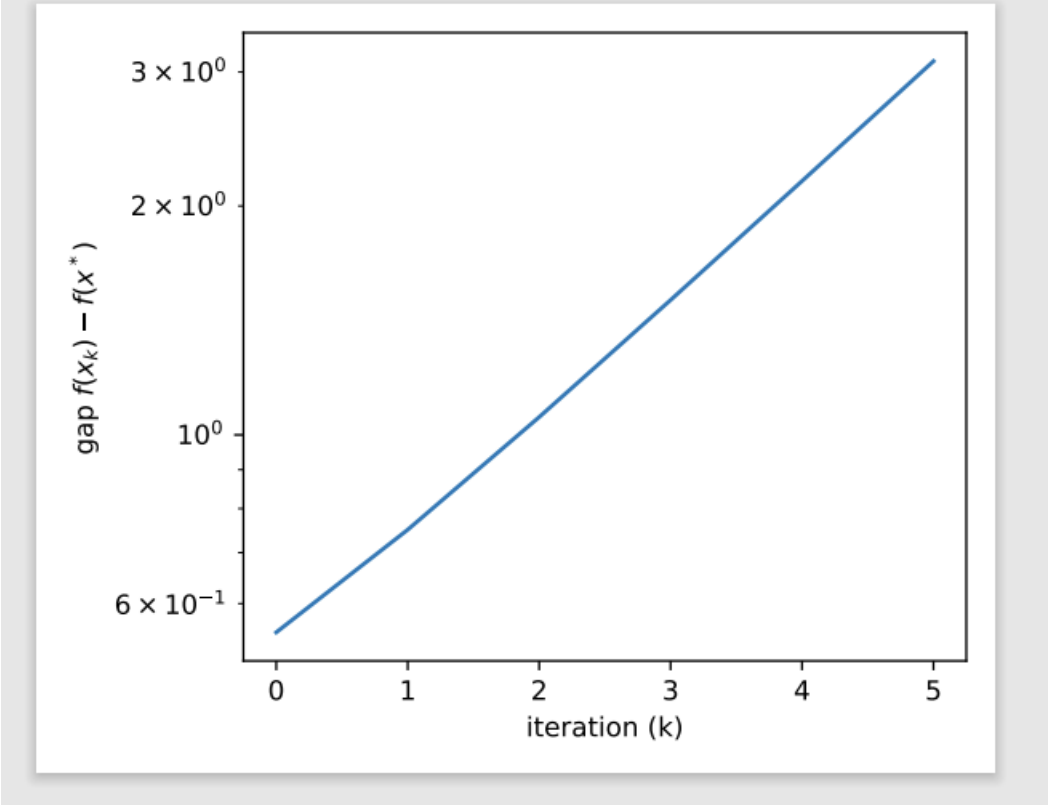
Ss = 1


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1


10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
gap $f(x_k) - f(x^*)$
0
20
40
60
80
iteration (k)

Ss = 2.2

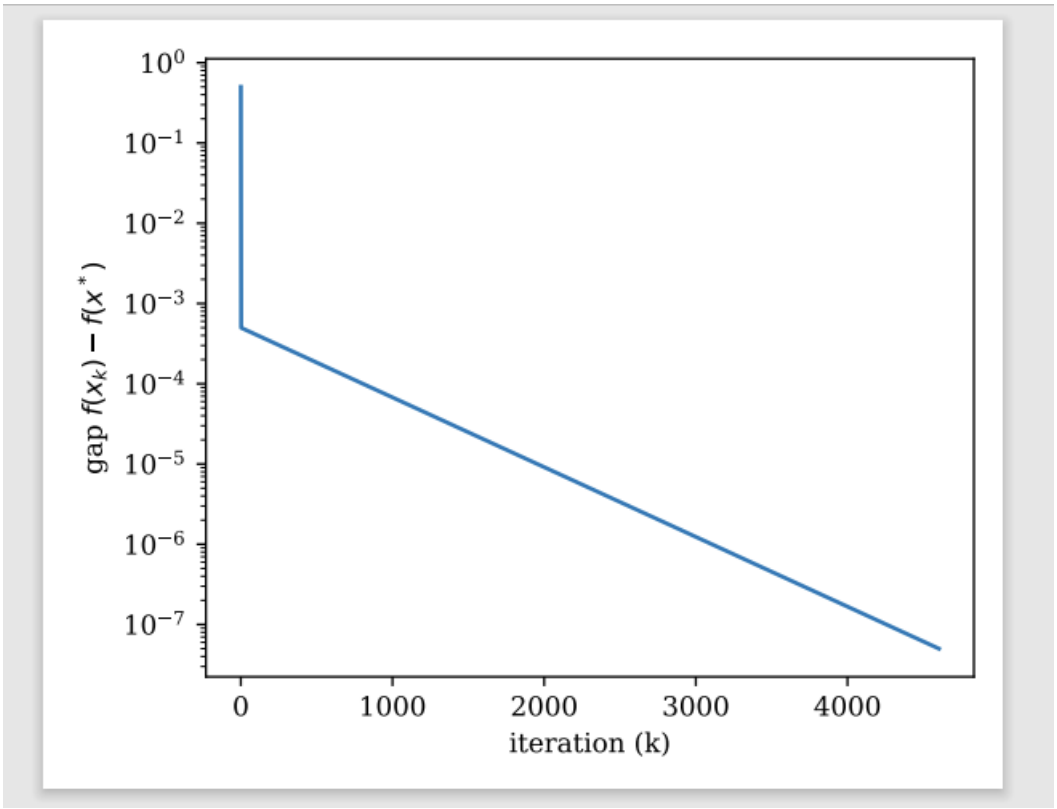

2
1
0
−1
−2
x2
0.2
0.4
0.6
0.8
1.0
x1

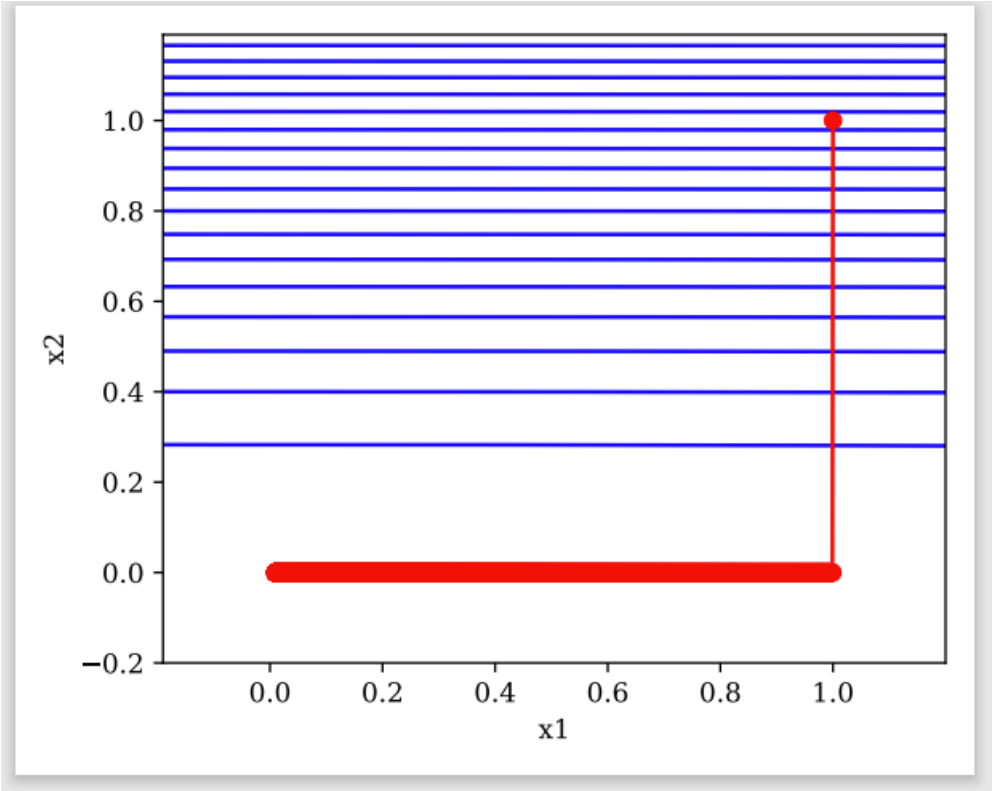


d)

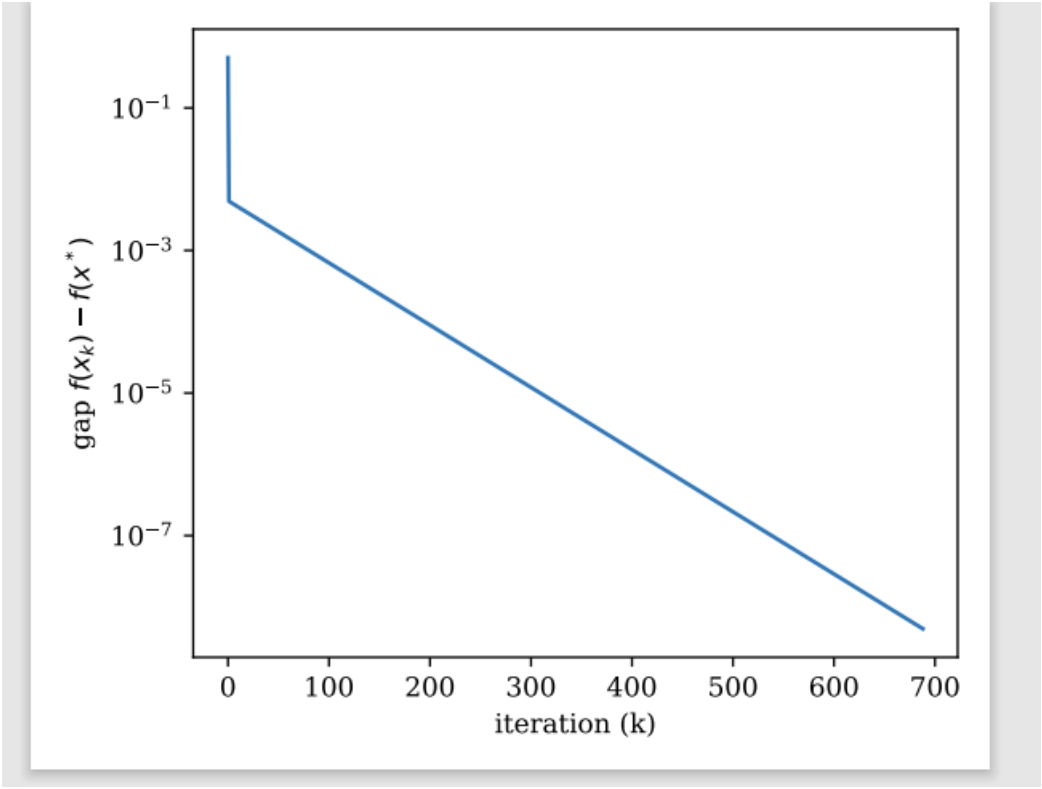
```
gamma=1, stepsize=1, number of iterations=1
gamma=0.1, stepsize=1, number of iterations=88
gamma=0.01, stepsize=1, number of iterations=688
gamma=0.001, stepsize=1, number of iterations=4603
```

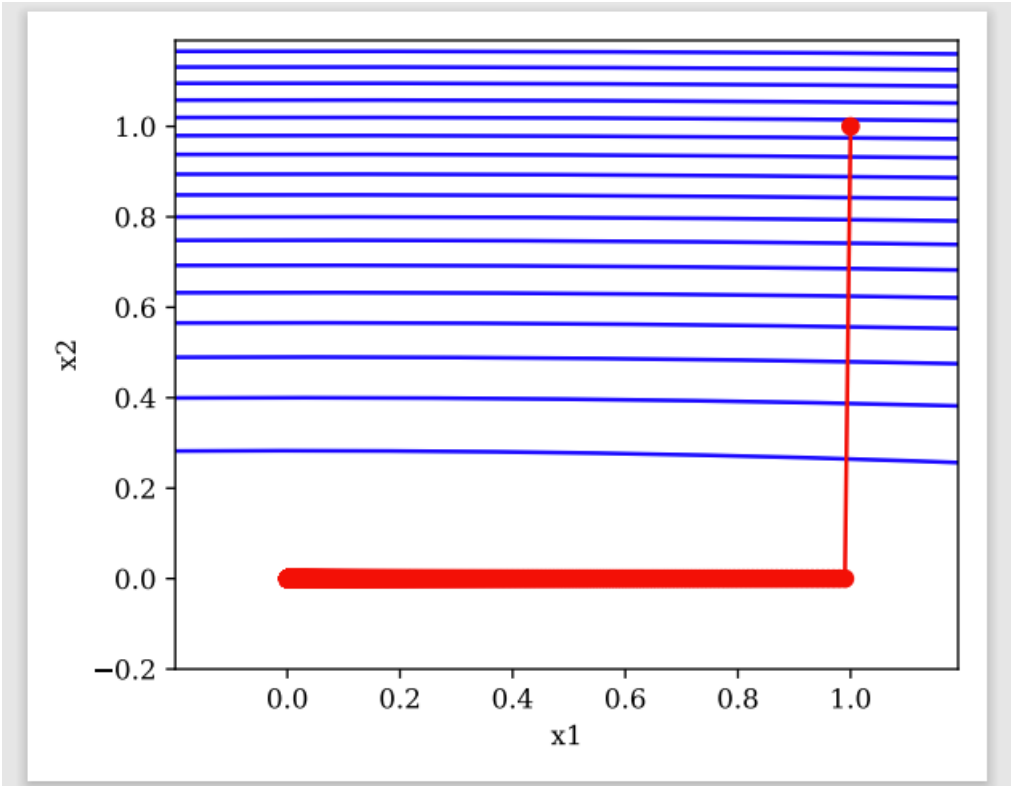
Gamma = 0.001


10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
gap $f(x_k) - f(x^*)$
0
1000
2000
3000
4000
iteration (k)

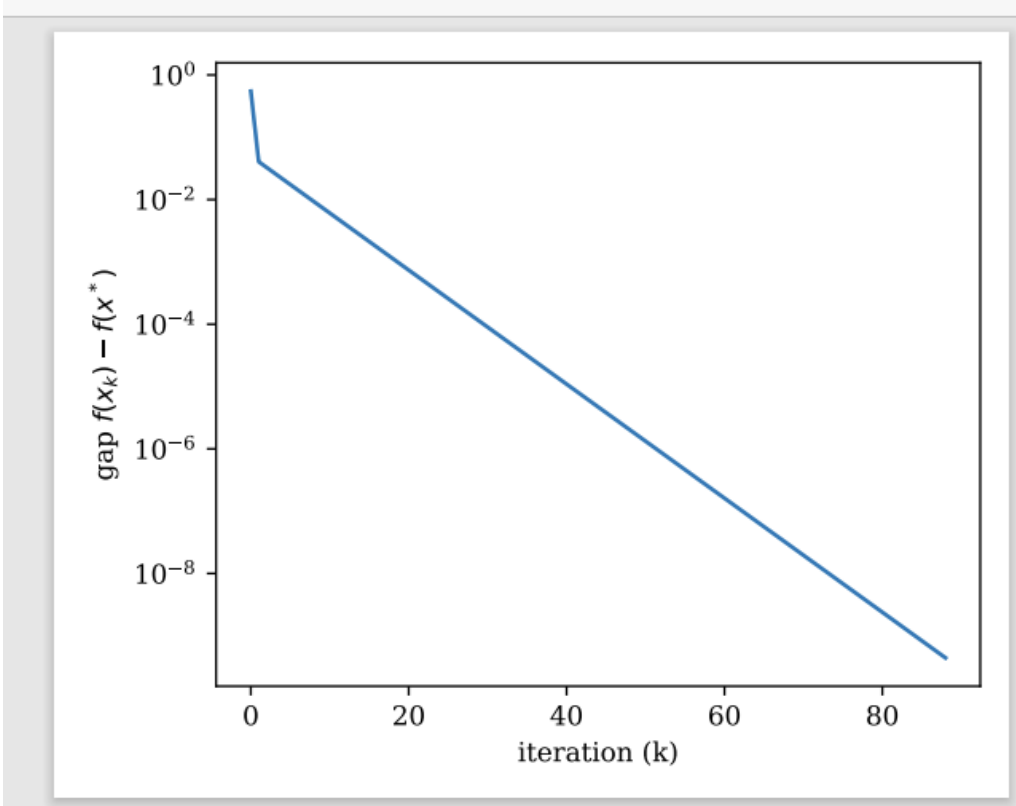

1.0
0.8
0.6
0.4
0.2
0.0
−0.2
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1

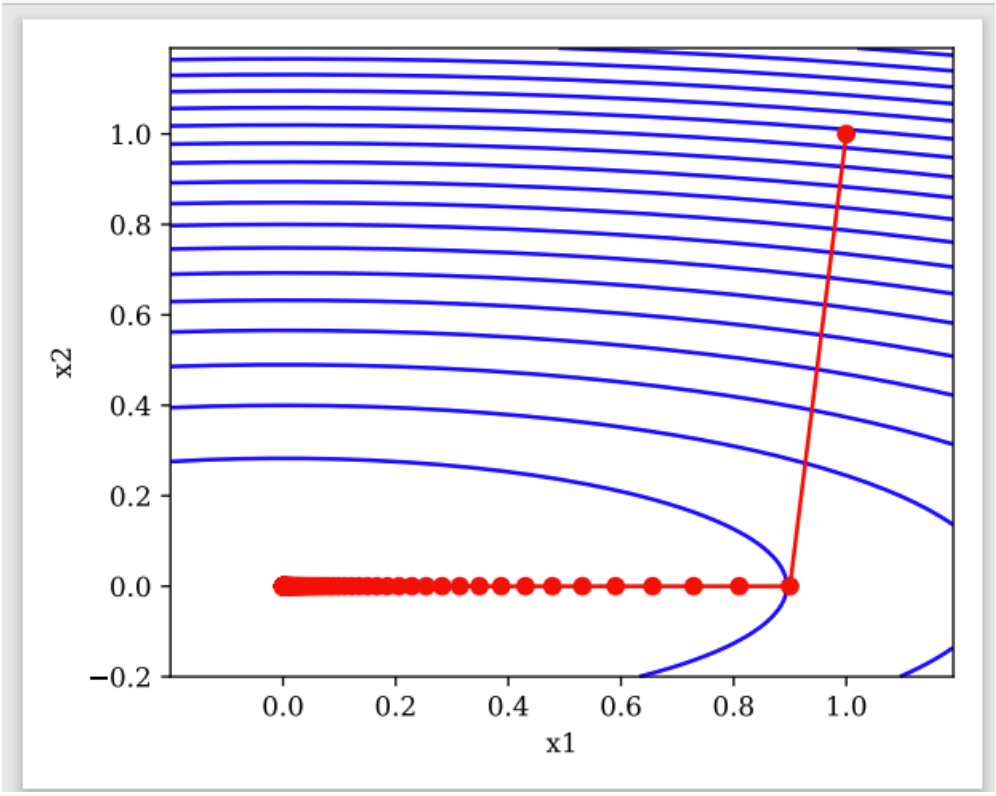
Gamma = 0.01


10^{-1}
10^{-3}
10^{-5}
10^{-7}
gap $f(x_k) - f(x^*)$
0
100
200
300
400
500
600
700
iteration (k)

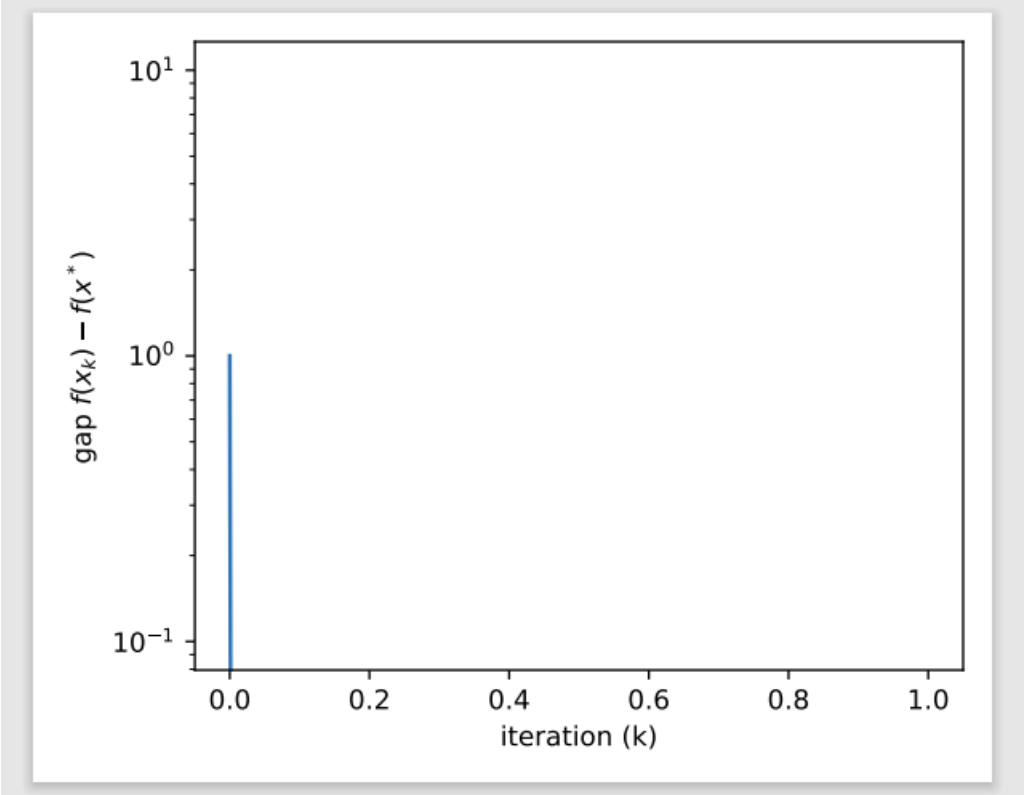

1.0
0.8
0.6
0.4
0.2
0.0
−0.2
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1

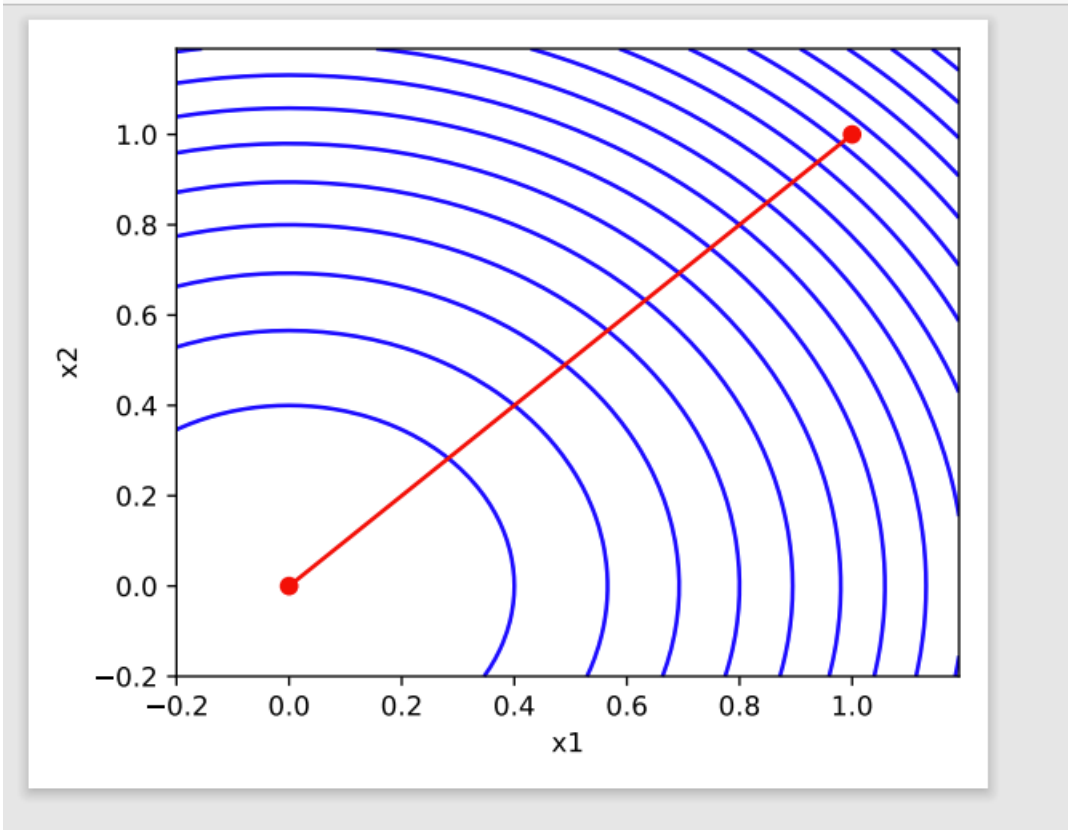
Gamma = 0.1


10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
gap $f(x_k) - f(x^*)$
0
20
40
60
80
iteration (k)


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1

Gamma = 1





2.

When using the same method as problem 1

```
stepsize=0.01, number of iterations=1146
optimal point [1.5        1.99999005]
```

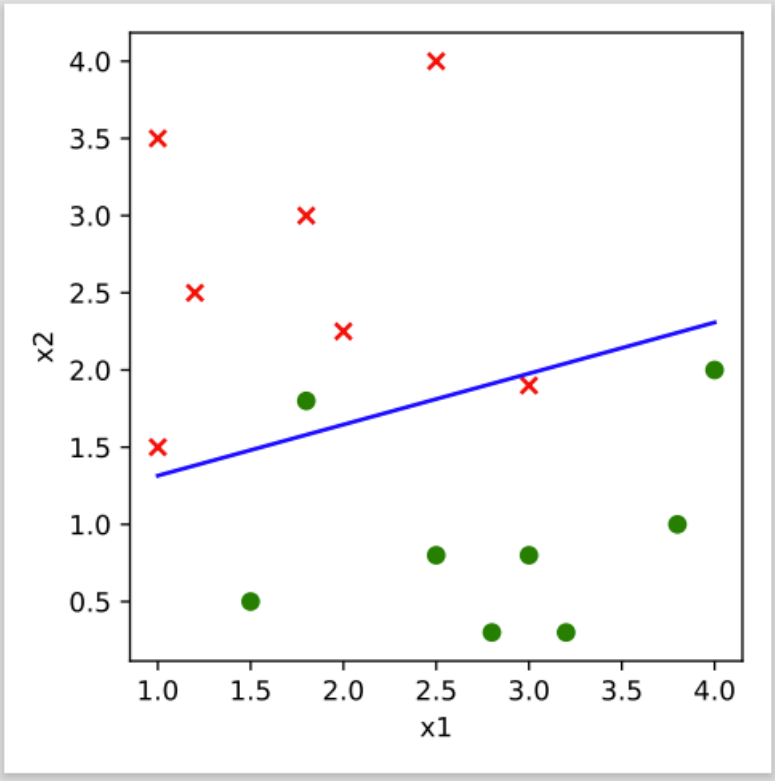
When solving the equation directly

```
solution [1.5 2. ]
```

Ignoring the errors, the answers agree with each other.

3.

```
[1.  1.5 1. ]
(15, 3) (15, 3) (15,)
accuracy = 0.8666666666666667
```



4. Because $g(x)$ is β-smooth, $-\beta I \leq \nabla^2 g(x) \leq \beta I$.

Because $f(x)$ is α-strongly convex, $f(x) - \frac{\alpha}{2}\|x\|^2$ is convex.

Namely, $\nabla^2\big(f(x)\big) \geq \alpha I$.

$\nabla^2 h(x) = \nabla^2\big(f(x)\big) - \nabla^2 g(x) \geq \alpha I - \beta I$.

If $\alpha \geq \beta$, $\nabla^2 h(x) \geq 0$.

So, $h(x)$ is convex.