SOLUTIONS HW 3

Problem 1

(a) Note that f has a unique global minimum at $x^* = 0$, $\nabla f(x) = 4x^3$, and $\nabla^2 f(x) = 12x^2$. Then for $x_k \neq 0$:

$$x_{k+1} = x_k - \frac{\alpha(4x_k)^3}{12x_k^2} = (1 - \frac{\alpha}{3})x_k.$$

Therefore, as long as $|1 - \frac{\alpha}{3}| < 1$, x_k converges to $x^* = 0$ as $k \to \infty$. The range of α can be found using $|1 - \frac{\alpha}{3}| < 1 \Rightarrow 0 < \alpha < 6$. Note that for $\alpha = 3$, the method converges in one step.

For this range of α and any $x_0 \in \mathbb{R}$, we can show

$$x_k = (1 - \frac{\alpha}{3})^k x_0,$$

hence x_k converges to 0 geometrically, i.e., the method converges "linearly".

(b)

$$\nabla f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \nabla^2 f(x) = \frac{4e^{2x}}{(e^{2x} + 1)^2}.$$

Substituting values in formula for Newton's method for $\alpha = 1$, we get the desired expression.

(For the extra numerical task, here is some example code:

```
import numpy as np
alpha = 1
x = 1
n = 5
iterates = np.zeros(n)
for i in range(n):
    x = x - (np.exp(4*x)-1)/(4*np.exp(2*x))
    iterates[i] = x
print(iterates)
```

For initialization $x_0 = 1$, iterates are:

$$[-8.13430204e-01 \quad 4.09402317e-01 \quad -4.73049165e-02 \quad 7.06028036e-05 \quad -2.34633642e-13].$$

For initialization $x_0 = 1.1$, iterates are:

$$[-1.12855259e+00 \quad 1.23413113e+00 \quad -1.69516598e+00 \quad 5.71536010e+00 \quad -2.30213565e+04].$$

The iterates converges to $x^* = 0$ with $x_0 = 1$ and diverges for $x_0 = 1.1$. We can see that Newton's method converges as long as the initial estimate is sufficiently close to x^*.)

(c) Since the cost function is quadratic and strongly convex, the Newton's method converges in one step as we discussed in the lecture note.

Problem 2

Since $f(x) = -\cos(x)$, we have $\nabla f(x) = \sin(x)$ and $\nabla^2 f(x) = \cos(x)$. The Newton's method with $\alpha = 1$ becomes: $x_{k+1} = x_k - \frac{\sin(x_k)}{\cos(x_k)} = x_k - \tan(x_k)$. Therefore, we have:

$$\begin{aligned} x_1 &= x_0 - \tan(x_0) \\ x_2 &= x_1 - \tan(x_1) = x_0 - \tan(x_0) - \tan(x_0 - \tan(x_0)). \end{aligned}$$

$x_2 = x_0 \Rightarrow \tan(-x_0) = \tan(x_0 - \tan(x_0))$. Therefore it suffices to show that $\exists\, x_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$ such that $-x_0 = x_0 - \tan(x_0) \Rightarrow 2x_0 = \tan(x_0)$. Let $h(x) = 2x - \tan(x), x \in [\frac{\pi}{4}, \frac{\pi}{2})$. Then $h(x)$ is continuous and $h(\frac{\pi}{4}) = \frac{\pi}{2} - 1 < 0$, $h(\frac{\pi}{4})$ tends to $+\infty$. By intermediate value theorem, there exists $x_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$ such that $h(x_0) = 0$. For this x_0, the Newton's methods does not converge since the iterates are oscillating between x_0 and x_1.

Problem 3

(a) Suppose to the contrary that there are two global minimizers x_1^* and x_2^*, where f takes the value of f^*. Then, strict convexity of f implies

$$f\left(\frac{x_1^* + x_2^*}{2}\right) < \frac{1}{2} f(x_1^*) + \frac{1}{2} f(x_2^*) = f^*$$

This is a contradiction as f attains a lower value at $\frac{x_1^*+x_2^*}{2}$ than f^*. Thus, f has a unique minimizer.

(b) One such example is $f(x) = e^{-x}$ and $\mathcal{S} = [0, \infty)$.

(c) All minimizers of f over $\mathcal{S}$ belong to nonempty sub-level sets of f intersected with $\mathcal{S}$. Next, we show that a sub-level set $\mathcal{X}_r := \{x | f(x) \leq r\}$ of f is compact for any r in the range of f. Consider any nonempty sublevel set $\mathcal{X}_r$. By the definition of the sub-level set and the continuity of f, this sub-level set is closed. Suppose the sub-level set is not bounded. Then, there is a sequence of $\{x_k\}$ in $\mathcal{X}_r$ such that $\|x_k\| \to \infty$. However, since f is radially unbounded, $f(x_k) \to \infty$, meaning that not all these points in the sequence can belong to the sublevel set. Thus, $\mathcal{X}_r$ is bounded. Hence, $\mathcal{X}_r$ is compact. Thus, $\mathcal{X}_r \cap S$ is compact, meaning that f must attain its minimum over this set. This minimum coincides with the minimum of f over $\mathcal{S}$.

(d) Consider a specific point $z \in \mathbb{R}$. Let $f(x) = \frac{1}{2} \|x - z\|^2$. Notice that z is fixed, while x varies over $\mathbb{R}^n$. Now, f is radially unbounded and is strictly convex. Part (c) shows that f is minimized uniquely over $\mathcal{S}$, meaning that projection is a well-defined function!

Problem 4

If $x \in [-1, 1]$, then the projection of x will be x. In this case, the distance from x to the projected point is 0, which cannot be further reduced (the distance has to be non-negative).

If $x > 1$, we need to minimize $|z - x|^2$ for $-1 \leq z \leq 1$. We can use our geometric insight to figure out that $|z - x|^2$ is monotonically decreasing for any $z \in (-\infty, x]$. Since $[-1, 1] \subset (-\infty, x]$, minimizing $|z - x|^2$ over $-1 \leq z \leq 1$ just requires setting $z = 1$. Hence the projected point is 1.

If $x < -1$, we can use a similar argument to figure out that the projected point is -1. (For fixed $x < -1$, the quadratic function $|z - x|^2$ is monotonically increasing for any $z \in [x, \infty)$.)