SOLUTIONS HW 4

1 Problem 1

The $\mathcal{S}$ is a closed convex set. The minimizer x^* is the projection of 0 in $\mathcal{S}$. Thus, in order to show that $x^* = A^T(AA^T)^{-1}b$ is the projection of 0 in $\mathcal{S}$ it suffices to show that

$$(x^* - 0)^T(x - x^*) \geq 0, \quad \forall\, x \in \mathcal{S} \tag{1}$$

Indeed,

$$\begin{aligned}
(x^*)^T(x - x^*) &= (b^T((AA^T)^{-1})^T A)(x - A^T(AA^T)^{-1}b) \\
&= b^T((AA^T)^{-1})^T Ax - b^T((AA^T)^{-1})^T AA^T(AA^T)^{-1}b \\
&= b^T((AA^T)^{-1})^T b - b^T((AA^T)^{-1})^T b, \quad \textit{we used } Ax{=}b \\
&= 0
\end{aligned} \tag{2}$$

2 Problem 2

1. Let us consider a vector x such that $x^T AA^T = 0$. Multiplying by x on the right, we have

$$x^T AA^T x = 0 \Rightarrow \|x^T A\|^2 = 0 \tag{3}$$

   Since the rows of A are linearly independent, we must have $x = 0$. Hence,

$$x^T AA^T = 0 \Rightarrow x = 0 \tag{4}$$

   which implies that AA^T is invertible.

2. In order to verify that $z^* = x - A^T(AA^T)^{-1}(Ax - b)$ is the project of x on $\mathcal{S}$ it suffices to show that $(z^* - x)^T(z - z^*) \geq 0$ for all $z \in \mathcal{S}$. Indeed,

$$\begin{aligned}
&(z^* - x)^T(z - z^*) \\
=&(x^T - (Ax - b)^T((AA^T)^{-1})^T A - x^T)(z - x + A^T(AA^T)^{-1}(Ax - b)) \\
=&(b - Ax)^T((AA^T)^{-1})^T Az + (Ax - b)^T((AA^T)^{-1})^T Ax - (Ax - b)^T((AA^T)^{-1})^T AA^T(AA^T)^{-1}(Ax - b) \\
=&(b - Ax)^T((AA^T)^{-1})^T b + (Ax - b)^T((AA^T)^{-1})^T Ax - (Ax - b)^T((AA^T)^{-1})^T Ax + (Ax - b)^T((AA^T)^{-1})^T b \\
=&0
\end{aligned} \tag{5}$$

3 Problem 3

1. The derivative of the Lagrangian is

$$\nabla f + \lambda \nabla h = 0 \Rightarrow 2x + \lambda \mathbf{1} = 0 \tag{6}$$

   This implies that $x_1 = \ldots = x_n = -\lambda/2$ and $\sum_{i=1}^n x_i = 2$. Hence, $x^* = [2/n, \ldots, 2/n]^T$.

   We can also check $\nabla^2_{xx} L(x^*, \lambda^*) = 2I \succ 0$. Hence x^* is a local min. Since f is coercive, we know the global min exists and the only local min x^* is also the global min.

2. The derivative of the Lagrangian is

$$\nabla f + \lambda \nabla h = 0 \Rightarrow \mathbf{1} + \lambda 2x = 0 \tag{7}$$

This implies that $x_1 = \ldots = x_n = -1/(2\lambda)$ and $\|x\|^2 = 1$. Hence, we have two stationary points $x^* = [1/\sqrt{n}, \ldots, 1/\sqrt{n}]^T$ or $x^* = -[1/\sqrt{n}, \ldots, 1/\sqrt{n}]^T$. However, for $x^* = -[1/\sqrt{n}, \ldots, 1/\sqrt{n}]^T$, we have

$$\nabla^2 f(x^*) + \lambda^* \nabla^2 h(x^*) = -\sqrt{n} I \prec 0$$

This is not a local min. For $x^* = -[1/\sqrt{n}, \ldots, 1/\sqrt{n}]^T$, we have

$$\nabla^2 f(x^*) + \lambda^* \nabla^2 h(x^*) = \sqrt{n} I \succ 0$$

This is a local min. Since the feasible set is compact, we know the global min exists and this point will also be the local min.