1　1

（1）

$$\begin{aligned}\mathrm{prox}_{th}(x) &= \arg\min_u th(u) + \frac{1}{2}\|u-x\|_2^2 \\ &= \arg\min_u \frac{t}{2}u^TAu + tb^Tu + tc + \frac{1}{2}(u-x)^T(u-x)\end{aligned}$$

令$f(u) = \frac{t}{2}u^TAu + tb^Tu + tc + \frac{1}{2}(u-x)^T(u-x)$，有

$$\nabla f(u) = tAu + tb^T + u - x$$

$\nabla f(u) = 0$的解为$u = (I+tA)^{-1}(x-tb)$，由于f是一个凸函数，这就是f的最小值点。

所以

$$\mathrm{prox}_{th}(x) = (I+tA)^{-1}(x-tb)$$

（2）

$$\begin{aligned}\mathrm{prox}_{th}(x) &= \arg\min_u th(u) + \frac{1}{2}\|u-x\|_2^2 \\ &= \arg\min_u \sum_{i=1}^n -t\log u_i + \frac{1}{2}(u_i-x_i)^2\end{aligned}$$

令$f(u) = \sum_{i=1}^n -t\log u_i + \frac{1}{2}(u_i-x_i)^2$，有

$$\nabla f(u)_i = -\frac{t}{u_i} + (u_i - x_i)$$

由求根公式知，$-\frac{t}{u_i} + (u_i-x_i)$的解为$u_i = \frac{x_i \pm \sqrt{x_i^2+4t}}{2}$，由于$t>0$并且log的定义域为$(0,+\infty]$，所以$\nabla f(u)=0$存在唯一解，

解为$u_i = \frac{x_i+\sqrt{x_i^2+4t}}{2}, \forall u_i$，由于$f$是一个凸函数，这就是$f$的最小值点。

所以

$$\mathrm{prox}_{th}(x)_i = \frac{x_i+\sqrt{x_i^2+4t}}{2}, i = 1,2,\cdots,n$$

（3）

$$x - u \in t\partial\|u\|_2 = \begin{cases}\left\{\frac{tu}{\|u\|_2}\right\}, & u \neq 0 \\ \{w : \|w\|_2 \leq t\}, & u = 0\end{cases}$$

先考虑$u \neq 0$的情况，$x - u = \frac{tu}{\|u\|_2}$，即$x = \left(\frac{t}{\|u\|_2}+1\right)u$，这说明$x$与$u$平行，也就是$\frac{u}{\|u\|_2} = \frac{x}{\|x\|_2}$，所以$u = x - \frac{tu}{\|u\|_2} = x - \frac{tx}{\|x\|_2} = \left(1-\frac{t}{\|x\|_2}\right)x$

对于$u=0$的情况，$x \in \{w : \|w\|_2 \leq t\}$，也就是$\|x\|_2 \leq t$

$$\mathrm{prox}_{th}(x) = \begin{cases}\left(1-\frac{t}{\|x\|_2}\right)x, & \|x\|_2 \geq t \\ 0, & \text{otherwise}\end{cases}$$

（4）

$$
\begin{aligned}
\mathrm{prox}_{th}(x) &= \arg\min{}_u th(u) + \frac{1}{2}\|u-x\|_2^2 \\
&= \arg\min{}_u \sum_{i=1}^{n} t|u_i| + \frac{1}{2}(u_i - x_i)^2
\end{aligned}
$$

令$f(u) = \sum_{i=1}^{n} t|u_i| + \frac{1}{2}(u_i-x_i)^2$，有

$$
\partial f(u)_i = \begin{cases} \{t + u_i - x_i\}, & u_i > 0 \\ [-t + u_i - x_i, t + u_i - x_i], & u_i = 0 \\ \{-t + u_i - x_i\}, & u_i < 0 \end{cases}
$$

于是

$$
\mathrm{prox}_{th}(x)_i = \begin{cases} x_i - t, & x_i \geq t \\ 0, & -t < x_i < t \\ x_i + t, & x_i \leq -t \end{cases}
$$

也就是

$$
\mathrm{prox}_{th}(x)_i = \mathrm{sign}(x)\max\{|x| - t, 0\}
$$

2　2

（1）

$$
\begin{aligned}
u_1 &= \mathrm{prox}_h(x) \\
&= \arg\min{}_u g(\lambda u + a) + \|u - x\|_2^2 \\
u_2 &= \mathrm{prox}_{\lambda^2 g}(\lambda x + a) \\
&= \arg\min{}_u \lambda^2 g(u) + \|u - \lambda x - a\|_2^2
\end{aligned}
$$

u_1, u_2需要满足最优条件(1)(2)

$$
x - u_1 \in \lambda \partial g(\lambda u_1 + a) \tag{1}
$$

$$
\lambda x - u_2 + a \in \lambda^2 \partial g(u_2) \tag{2}
$$

将$u_1 = \frac{1}{\lambda}(u_2 - a)$带入(1)

$$
x - \frac{1}{\lambda}(u_2 - a) \in \lambda \partial g(u_2) \tag{3}
$$

由于$\lambda \neq 0$，(3)与(2)完全等价，所以$u_1 = \frac{1}{\lambda}(u_2 - a)$正确描述了$u_1, u_2$之间的关系，也就是

$$
\mathrm{prox}_h(x) = \frac{1}{\lambda}(\mathrm{prox}_{\lambda^2 g}(\lambda x + a) - a)
$$

（2）

$$
\begin{aligned}
u_1 &= \mathrm{prox}_h(x) \\
&= \arg\min{}_u \lambda g\left(\frac{u}{\lambda}\right) + \|u - x\|_2^2 \\
u_2 &= \mathrm{prox}_{\lambda^{-1} g}\left(\frac{x}{\lambda}\right) \\
&= \arg\min{}_u \lambda^{-1} g(u) + \left\|u - \frac{x}{\lambda}\right\|_2^2
\end{aligned}
$$

u_1, u_2需要满足最优条件(1)(2)

$$
x - u_1 \in \partial g\left(\frac{u_1}{\lambda}\right) \tag{1}
$$

$$
\frac{x}{\lambda} - u_2 \in \lambda^{-1} \partial g\left(u_2\right) \tag{2}
$$

将$u_1 = \lambda u_2$带入(1)

$$
x - \lambda u_2 \in \partial g(u_2) \tag{3}
$$

$$x-\lambda u_2 \in \partial g(u_2) \tag{3}$$

由于$\lambda \neq 0$，(3)与(2)完全等价，所以$u_1=\lambda u_2$正确描述了u_1, u_2之间的关系，也就是

$$\operatorname{prox}_h(x)=\lambda \operatorname{prox}_{\lambda^{-1} g}\left(\frac{x}{\lambda}\right)$$

（3）

$$\begin{aligned}
u_1 & =\operatorname{prox}_h(x) \\
& =\arg \min _u g(u)+a^T u+\|u-x\|_2^2 \\
u_2 & =\operatorname{prox}_g(x-a) \\
& =\arg \min _u g(u)+\|u-x+a\|_2^2
\end{aligned}$$

u_1, u_2需要满足最优条件(1)(2)

$$x-u_1-a \in \partial g(u_1) \tag{1}$$

$$x-u_2-a \in \partial g(u_2) \tag{2}$$

(1)与(2)完全等价，所以

$$\operatorname{prox}_h(x)=\operatorname{prox}_g(x-a)$$

3　3

(1) 对于$y \in \mathbb{R}^n$，y在C上的投影就是下面这个优化问题的最优解

$$\begin{aligned}
& \min \|x-y\|_2^2 \\
& \text{s.t.} \quad Ax=b
\end{aligned}$$

上述约束优化问题的拉格朗日函数为

$$L=\|x-y\|_2^2+v^T A x$$

该函数可微

$$\begin{aligned}
\nabla L=2x-2y+A^T v=0 \\
2Ax-2Ay+AA^T v=0
\end{aligned}$$

于是可以分别解出v, x

$$\begin{gathered}
v=2(AA^T)^{-1}(b-Ay) \\
x=y-A^T v=y-A^T(AA^T)^{-1}(b-Ay)
\end{gathered}$$

x就是y在C上的投影，所以

$$P_C(x)=x-A^T(AA^T)^{-1}(b-Ax)$$

（2）对于$y \in \mathbb{R}^n$，y在C上的投影就是下面这个优化问题的最优解

$$\begin{aligned}
& \min \|x-y\|_2^2 \\
& \text{s.t.} \quad \|x\|_1 \leq 1
\end{aligned}$$

对于上述优化问题，如果$\|y\|_1 \leq 1$，其最优解为$x=y$。

如果$\|y\|_1>1$，为方便求解，我们取$f(x)=\frac{1}{2}\|x-y\|_2^2$，写出拉格朗日函数

$$L=\frac{1}{2}\|x-y\|_2^2+w(\|x\|_1-1)$$

L每个分量的次梯度表示为

$$\partial L_i = \begin{cases} \{x_i - y_i + w\}, & x_i > 0 \\ \{x_i - y_i + tw | -1 \le t \le 1\}, & x_i = 0 \\ \{x_i - y_i - w\}, & x_i < 0 \end{cases}$$

最优解需要满足稳定性条件$0 \in L_i$。对于$x_i > 0$的情况，$x_i = y_i - w > 0$，$y_i > w$；对于$x_i = 0$的情况，$y_i = tw \in [-w, w]$；对于$x_i < 0$情况，$x_i = y_i + w < 0$，$y_i < -w$。综合三种情况，可以写为$x_i = \max\{|y_k| - w, 0\}$。对于约束条件$\|x\|_1 = 1$，即为$\sum_{i=1}^{n}\{|y_k| - w, 0\} = 0$。

所以x到l_1范数的单位球的投影为

$$P_C(x)_k = \begin{cases} x_k - \lambda, & x_k > \lambda \\ 0, & -\lambda \le x_k \le \lambda \\ x_k + \lambda, & x_k < -\lambda \end{cases}$$

如果$\|x\|_1 \le 1$，$\lambda = 0$；否则λ满足$\sum_{i=1}^{n}\{|x_k| - \lambda, 0\} = 0$。