HOMEWORK 5

QUESTION 1

According to the first-order optimality condition, the solution to the following problem:

$$\min_{\boldsymbol{x}\in\bar{B}} f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_0\|^2 \tag{1}$$

Should satisfy:

$$2(\boldsymbol{x}^*-\boldsymbol{x}_0)^T\cdot(\boldsymbol{x}-\boldsymbol{x}^*)\geq 0, \forall \boldsymbol{x}\in\bar{B} \tag{2}$$

Where $\boldsymbol{x}^*$ is the projection of $\boldsymbol{x_0}$ onto $\bar{B}$.

First, we show that $\boldsymbol{x}^*\in\partial B$. Otherwise, $\exists\epsilon>0, \exists\boldsymbol{x}\in O(\boldsymbol{x}^*,\epsilon), s.t.\ (\boldsymbol{x}-\boldsymbol{x}^*)$ and $(\boldsymbol{x}^*-\boldsymbol{x_0})$ are collinear but reversed, thus contradicting condition (2).

Second, we show that $\boldsymbol{x}^*$ and $\boldsymbol{x_0}$ are collinear. Otherwise, we could make $\boldsymbol{x}$ close enough to $\boldsymbol{x}^*$ to make $(\boldsymbol{x}-\boldsymbol{x}^*)$ lie in the direction of the tangent line of $\bar{B}$ at $\boldsymbol{x}^*$, which lead to $2(\boldsymbol{x}^*-\boldsymbol{x}_0)^T\cdot(\boldsymbol{x}-\boldsymbol{x}^*)\leq 0$.

Combining the two conclusions above, we derive that $\hat{x_0}=\frac{x_0}{||x_0||}$

QUESTION 2

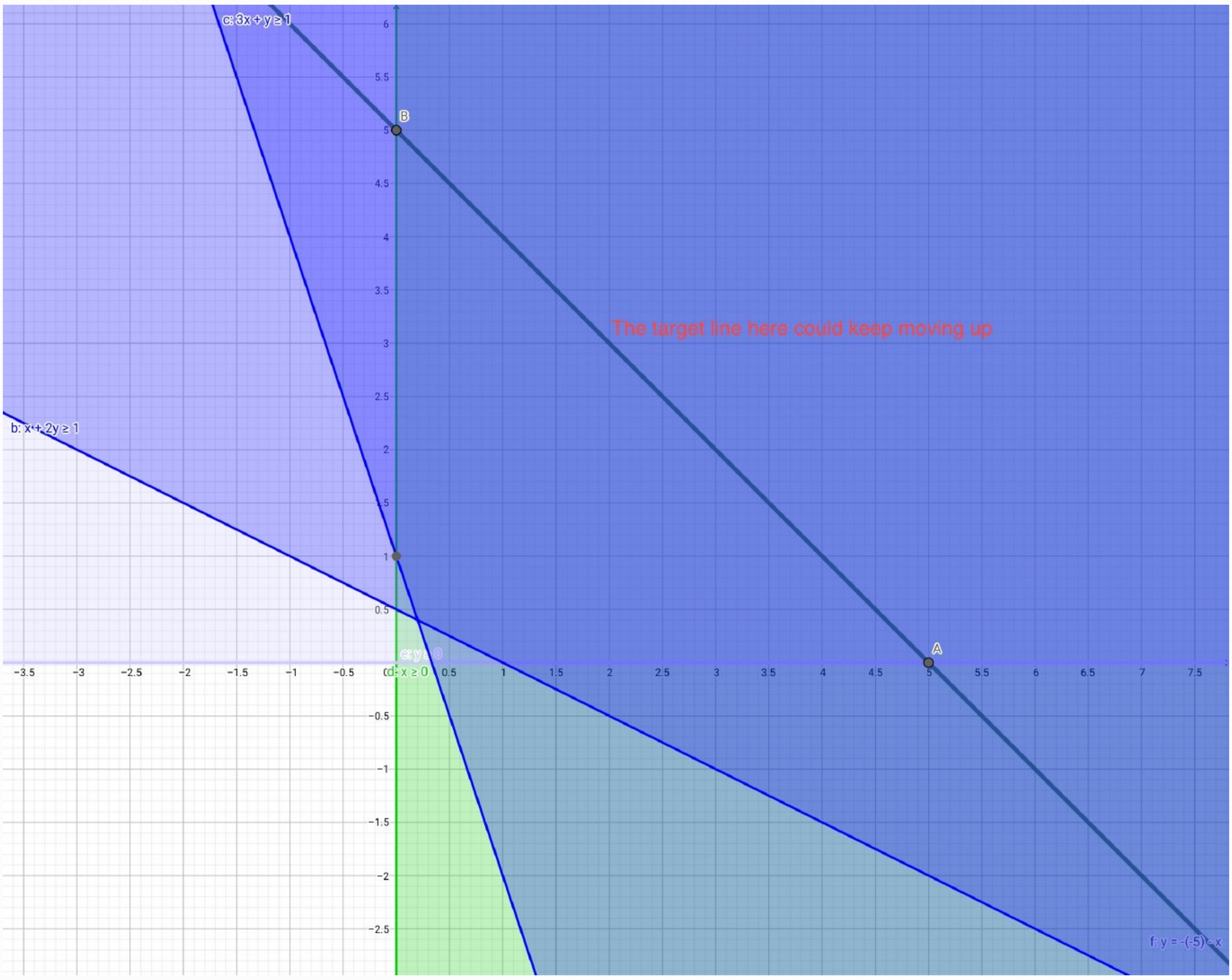
(a).

```
status: optimal
optimal value: 0.5999999999116254
optimal var: x1=0.1999999999391762 x2=0.3999999999724492
```

Optimal solution: $x_1 = 0.2, x_2 = 0.4$

Optimal value: 0.6

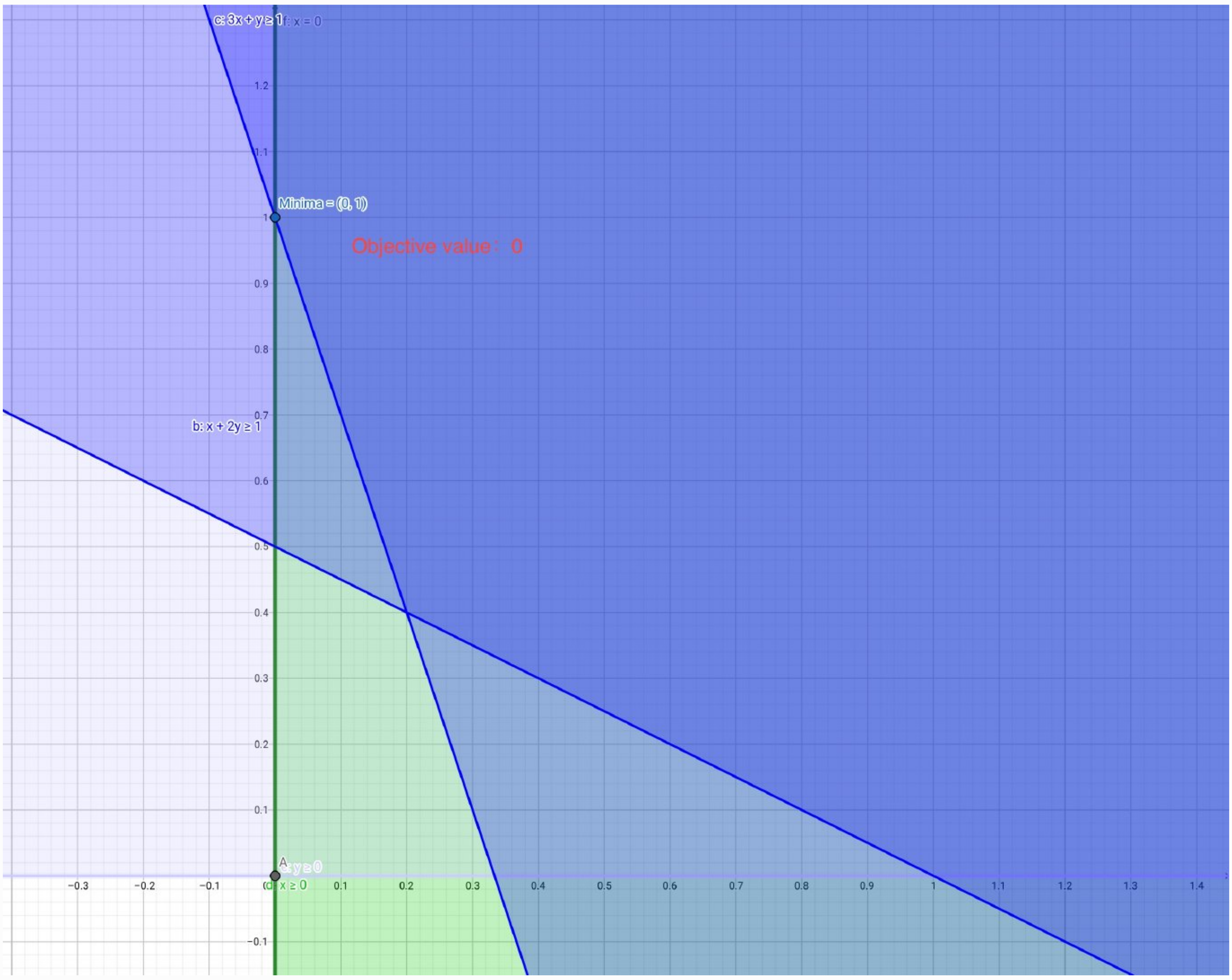
(b).




```
status: unbounded
optimal value: -inf
optimal var: x1=None x2=None
```

No No optimal solution. Or the Optimal is $+\infty$, it's infeasible.

(c).




```
status: optimal
optimal value: -1.232214801046685e-10
optimal var: x1=-1.232214801046685e-10 x2=1.7673174212389093
```

Optimal solution: $\{(x_1, x_2)|x_1 = 0, x_2 \geq 1\}$

Optimal value: 0

(d).

```
status: optimal
optimal value: 0.3333333334080862
optimal var: x1=0.3333333286259564 x2=0.3333333334080862
```

(e).

```
status: optimal
optimal value: 0.6923076924267746
optimal var: x1=0.6923065264125655 x2=0.1538467368351237
```

PROBLEM 3

(a).

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_\infty \\ \text{s.t.} \quad & \|\boldsymbol{x}\|_\infty \leq 1 \end{aligned} \tag{3}$$

We introduce a new variable s:

$$
\begin{aligned}
\min_{x_1,x_2,\cdots,x_n,s} \quad & s \\
\text{s.t.} \quad & -1 \le x_i \le 1, i = 1,2,3,\cdots n \\
& -s \le \boldsymbol{a}_i^T x_i - b_i \le s, i = 1,2,3,\cdots,m
\end{aligned}
\tag{4}
$$

Here $\boldsymbol{a_i}$ is the i-row vector of $\boldsymbol{A}$, b_i is the i-component of $\boldsymbol{b}$.

For conciseness, it can be further interpreted as:

$$
\begin{aligned}
\min_{\boldsymbol{x},s} \quad & s \\
\text{s.t.} \quad & -\boldsymbol{1} \le \boldsymbol{x} \le \boldsymbol{1} \\
& -\boldsymbol{s} \le \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \le \boldsymbol{s}
\end{aligned}
\tag{5}
$$

Here $\boldsymbol{s}$ means the vector whose components are all s.

(b).

```
status: optimal
optimal value: 5.333333333553781
optimal var: x = [-0.33333333  0.33333333]
```

(c).

```
status: optimal
optimal value: 5.333333333260567
optimal var: x = [-0.33333333  0.33333333], s = 5.333333333260567
```

PROBLEM 4

(a.)

The normal equation is:

$$
\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{w} = \boldsymbol{X}^T\boldsymbol{y} \tag{6}
$$

We use PYTHON Numpy to solve the equation.

We first notice that $\boldsymbol{X}$ is invertible, so we get:

$$
\boldsymbol{w} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{7}
$$

```
[ 1.22170662 -0.21469307  0.15549204 -0.4586777   1.18537706  0.00613317]
```

(b).

t = 1

```
status: optimal
optimal value: 31.314550054478023
optimal var:[5.54241960e-01 4.31525539e-09 9.92071629e-10 9.38255329e-09
 4.30602870e-01 1.51551568e-02]
```

Compared to the result in (a). **the solution isn't the same.**

Here 3 out of 6 components are equal to zero. **Thus it is a sparse solution.**

t = 10

```
status: optimal
optimal value: 13.295569218508426
optimal var:[ 1.22171615 -0.21469843  0.15549443 -0.45868521  1.18537859  0.00613412]
```

Compared to the result in (a). **the solution is almost the same.**

Here only the last component could be considered zero.

Thus it isn't a sparse solution.

(c).

t = 1

```
status: optimal
optimal value: 16.173131057359125
optimal var:[0.52516383 0.08616926 0.09403005 0.12515129 0.82965381 0.06283205]
```

Compared to the result in (a). **the solution isn't the same.**

It has no zero component.

t = 100

```
status: optimal
optimal value: 13.295569218196668
optimal var:[ 1.22170662 -0.21469308  0.15549205 -0.4586777   1.18537705  0.00613318]
```

Compared to the result in (a). **the solution is almost the same.**

It has no zero component.

Summary for Problem 4

From the results above, we verify that `Lasso` tends to produce a sparse solution with zero components in the Linear least squares regression problem, compared to `Ridge regression`.