# Linear and Convex Optimization Homework 05

Qiu Yihang, 2021/11/01-11/04

1. ***Proof:***

Let $\hat{\boldsymbol{x}}_{\boldsymbol{0}}$ be the solution to the following optimization problem

$$\min_{\boldsymbol{x}\in\bar{B}} f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_{\boldsymbol{0}}\|^2.$$

By the definition of the projection onto a convex set, $\hat{\boldsymbol{x}}_0$ is the projection of $\boldsymbol{x}_0$ onto $\bar{B}$.

By the first-order optimality condition, we know $\nabla f(\hat{\boldsymbol{x}}_0)(\boldsymbol{x}-\hat{\boldsymbol{x}}_0) \geq 0, \forall \boldsymbol{x}\in\bar{B}$.

$$\nabla f(\boldsymbol{x}) = \nabla\left(\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^T(\boldsymbol{x}-\boldsymbol{x}_0)\right) = (\boldsymbol{x}-\boldsymbol{x}_0)^T$$

Thus, $(\hat{\boldsymbol{x}}_0-\boldsymbol{x}_0)^T(\boldsymbol{x}-\hat{\boldsymbol{x}}_0) \geq 0, \forall \boldsymbol{x}\in\bar{B}.$

Now we prove $\hat{\boldsymbol{x}}_0 = \frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}$ by contradiction.

Let $\bar{\boldsymbol{x}}_0 = \frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|} \in \partial\bar{B}$. Assume $\hat{\boldsymbol{x}}_0 \neq \bar{\boldsymbol{x}}_0$. Since $\hat{\boldsymbol{x}}_0\in\bar{B}$, $\|\hat{\boldsymbol{x}}_0\| \leq 1$.

Obviously $\hat{\boldsymbol{x}}_0^T\bar{\boldsymbol{x}}_0 = \bar{\boldsymbol{x}}_0^T\hat{\boldsymbol{x}}_0 \leq \|\hat{\boldsymbol{x}}_0\|\|\bar{\boldsymbol{x}}_0\| \leq 1.$

Therefore,

$$\begin{aligned}(\hat{\boldsymbol{x}}_0-\boldsymbol{x}_0)^T(\bar{\boldsymbol{x}}_0-\hat{\boldsymbol{x}}_0) &= \hat{\boldsymbol{x}}_0^T\bar{\boldsymbol{x}}_0 - \boldsymbol{x}_0^T\bar{\boldsymbol{x}}_0 - \hat{\boldsymbol{x}}_0^T\hat{\boldsymbol{x}}_0 + \boldsymbol{x}_0^T\hat{\boldsymbol{x}}_0 = (1+\|\boldsymbol{x}_0\|)\bar{\boldsymbol{x}}_0^T\hat{\boldsymbol{x}}_0 - \|\boldsymbol{x}_0\| - 1\\ &= (1+\|\boldsymbol{x}_0\|)(\bar{\boldsymbol{x}}_0^T\hat{\boldsymbol{x}}_0 - \mathbf{1}) \leq 0.\end{aligned}$$

Meanwhile, $(\hat{\boldsymbol{x}}_0-\boldsymbol{x}_0)^T(\boldsymbol{x}-\hat{\boldsymbol{x}}_0) \geq 0, \forall \boldsymbol{x}\in\bar{B}$. **Contradiction!**

In conclusion, $\hat{\boldsymbol{x}}_0 = \frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}$.

*Qed.* ■

2. ***Solution:***

First, we draw the visualization of five optimization problems and try to find the set of optimal solutions and the optimal value graphically.

Rough visualization of five optimization problems are as follows (see Fig.01).

**(a)** From the visualization, we know the minimum value is $\frac{3}{5}$ at $\left(\frac{2}{5},\frac{1}{5}\right)$.

**(b)** From the visualization, we know the objective function is unbounded.

**(c)** From the visualization, we know the minimum value is $0$ at $(0,y), \forall y \geq 1$.

**(d)** From the visualization, we know the minimum value is $\frac{1}{3}$ at $\left(\frac{1}{3},\frac{1}{3}\right)$.

**(e)** From the visualization, we know the minimum value is $\frac{1}{2}$ at $\left(\frac{1}{2},\frac{1}{6}\right)$.

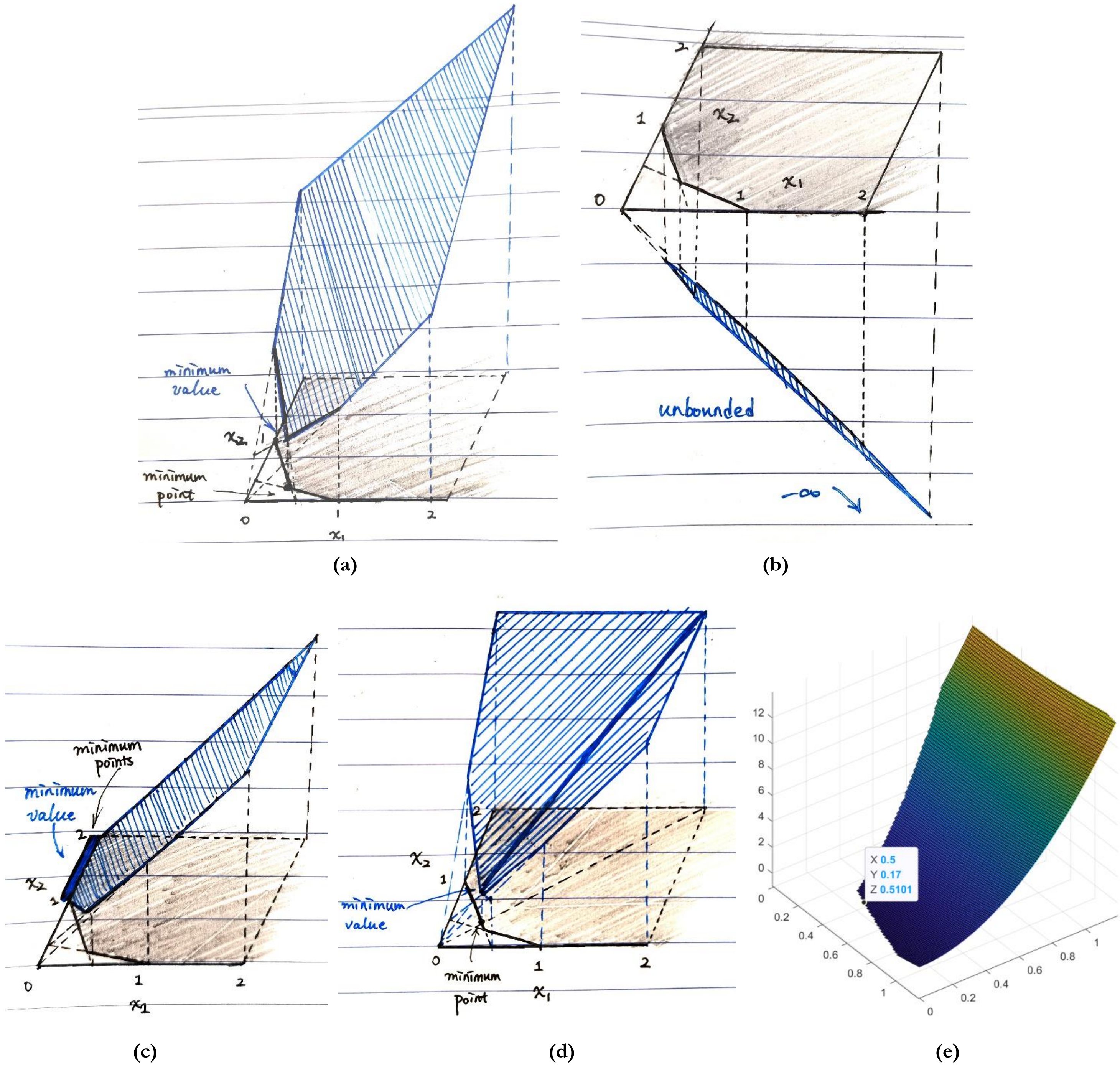


Fig.01 Visualization of Optimization Problems in Question 2

Use CVXPY to solve the problems given. The code is given in the zip file with name "110302.py". The results are given below (shown in Fig.02).

```
In [1]: runfile('D:/Textbooks/2021-2022-1/Linear and Convex Optimization/110302.py',
wdir='D:/Textbooks/2021-2022-1/Linear and Convex Optimization')
Question 2(a)
status:  optimal
optimum value:  0.5999999999116253
optimum var: [ 0.3999999999724491 , 0.1999999999391762 ]
```

```
Question 2(b)
status:  unbounded
optimum value:  -inf
optimum var: [ None , None ]

Question 2(c)
status:  optimal
optimum value:  -2.2491441767693299e-10
optimum var: [ -2.2491441767693299e-10 , 1.5537158969947242 ]

Question 2(d)
status:  optimal
optimum value:  0.3333333334080862
optimum var: [ 0.3333333334080862 , 0.33333333286259564 ]

Question 2(e)
status:  optimal
optimum value:  0.5000000000000002
optimum var: [ 0.5000000000000001 , 0.1666666666666667 ]
```

Fig 02. Results of Program 2

**Reference:** the first example at https://www.cvxpy.org/tutorial/intro/index.html.

Check the results. We find that the results are approximate to our calculations by hand, each with a numerical error at $10^{-10}$ level or even smaller. ■

3. ***Solution:***

**(a)** The optimization problem can be reformulated as

$$\min_{\boldsymbol{x},\boldsymbol{y}} \boldsymbol{1}^T\boldsymbol{y}$$

$$\begin{aligned} \text{s.t. } & \boldsymbol{x} \leq \boldsymbol{1}, \\ & -\boldsymbol{x} \leq \boldsymbol{1}, \\ & -\boldsymbol{y} \leq -\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}, \\ & -\boldsymbol{y} \leq \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \end{aligned}$$ ■

**(b)** Use CVXPY to solve the problem. The code is given in the zip file with name "110303.py". The result is as follows (shown in Fig.03). ■

```
In [3]: runfile('D:/Textbooks/2021-2022-1/Linear and Convex Optimization/110303.py',
wdir='D:/Textbooks/2021-2022-1/Linear and Convex Optimization')
Question 3(b)
status:  optimal
optimum value:  13.999999990735517
optimum var:
[ [[ 1.]
 [-1.]] ]
```

Fig 03. Results of Program 3(b)

**Reference:** https://www.cvxpy.org/tutorial/intro/index.html#vectors-and-matrices.

**(c)** Use CVXPY to solve the problem. The code is given in the zip file with name "110303.py". The result is as follows (shown in Fig.04). ■

```
Question 3(c)
status:  optimal
optimum value:  13.999999998611603
optimum var:
[ [[ 1.]
 [-1.]] ]
```

Fig 04. Results of Program 3(c)

**4.(a)** ***Solution:***

The normal equation of the problem is $\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{w}-\boldsymbol{y})=\boldsymbol{0}$ **i.e.** $\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{w}=\boldsymbol{X}^T\boldsymbol{y}$ with $\boldsymbol{X}^T\boldsymbol{y}=\begin{pmatrix}6\\2\end{pmatrix}$ and $\boldsymbol{X}^T\boldsymbol{X}=\begin{pmatrix}4&0\\0&1\end{pmatrix}$.

Since $\boldsymbol{X}$ has full column rank, we know the unique solution is

$$\boldsymbol{w}^\star=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}=\begin{pmatrix}1.5\\2\end{pmatrix}. \qquad ■$$

The minimum value is $\min_{\boldsymbol{w}}\|\boldsymbol{X}\boldsymbol{w}-\boldsymbol{y}\|_2^2=\|\boldsymbol{X}\boldsymbol{w}^\star-\boldsymbol{y}\|_2^2=4$. ■

**(b)** ***Solution*:**

Use CVXPY to solve the problem. The code is given in the zip file with name "110304.py". The result is given below (shown in Fig.05).

```
In [2]: runfile('D:/Textbooks/2021-2022-1/Linear and Convex Optimization/110304.py',
wdir='D:/Textbooks/2021-2022-1/Linear and Convex Optimization')
Question 4(b)
Lasso with t = 1 :
status:  optimal
optimum value:  9.0000000633334
optimum var:
[ [[9.99962136e-01]
 [3.78500362e-05]] ]

Lasso with t = 10 :
status:  optimal
optimum value:  4.000000000012024
optimum var:
[ [[1.49999883]
 [1.99999744]] ]
```

Fig. 05. Results of Program 4(b)

The result of lasso with $t=1$ is not the same as that of (a). Ignoring numerical errors, the solution should be $\begin{pmatrix}1\\0\end{pmatrix}$ and thus has a zero component.

The result of lasso with $t=10$ is exactly the same as that of (a) (ignoring numerical errors). The solution has no zero components. ■

**(c)** ***Solution*:**

Use CVXPY to solve the problem. The code is given in the zip file with name "110304.py". The result is given below (shown in Fig.06).

```
Question 4(c)
Ridge Regression with t = 1 :
status:  optimal
optimum value:  7.857489699180057
optimum var:
[ [[0.86266947]
 [0.50576813]] ]

Ridge Regression with t = 100 :
status:  optimal
optimum value:  4.0
optimum var:
[ [[1.5]
 [2. ]] ]
```

Fig. 06. Results of Program 4(c)

The result of ridge regression with $t = 1$ is not the same as that of (a). The solution has no zero components.

The result of ridge regression with $t = 100$ is exactly the same as that of (a). The solution has no zero components. ■