SUN Yilin
520030910361

Homework 5

1

By the first-order optimality condition we know that the solution $\boldsymbol{x}^*$ to the given problem must satisfy that $\nabla f(\boldsymbol{x}^*)^T(\boldsymbol{x}-\boldsymbol{x}^*) \geq 0, \forall \boldsymbol{x} \in \bar{B}$. By taking the gradient of f we know that the inequality is equivalent to $(\boldsymbol{x}^*-\boldsymbol{x}_0)^T(\boldsymbol{x}-\boldsymbol{x}^*) \geq 0$. We already know that this inequality is also the condition of the projection of point $\boldsymbol{x}_0$ onto $\bar{B}$. Now to show that the projection of $\boldsymbol{x}_0 \notin \bar{B}$ is $\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}$, we only need to show that $\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}$ satisfies $(\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}-\boldsymbol{x}_0)^T(\boldsymbol{x}-\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}) \geq 0$. Then by transformation, that is equivalent to show $\boldsymbol{x}_0^T\boldsymbol{x} \leq ||\boldsymbol{x}_0||$. Then by applying $H\ddot{o}lder's$ $Inequality$ we get $\boldsymbol{x}_0^T\boldsymbol{x} = \sum_{i=1}^n x_{0_i}x_i \leq \sum_{i=1}^n |x_{0_i}x_i| \leq ||\boldsymbol{x}_0|| \cdot ||\boldsymbol{x}|| \leq ||\boldsymbol{x}_0||$. So $\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}$ does satisfy $(\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}-\boldsymbol{x}_0)^T(\boldsymbol{x}-\frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||}) \geq 0$. So it is the projection of $\boldsymbol{x}_0 \notin \bar{B}$.

2

Below is the feasible set of the problem. Note that the area is directed by the arrows.

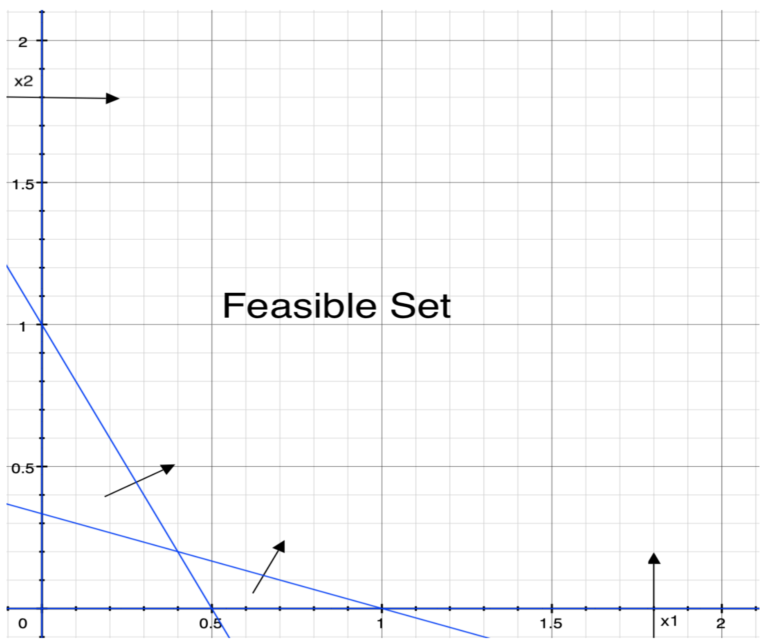

Figure 1: Feasible Set for Problem 2

(a)

Graphically, the set of optimal solutions is a single point, i.e. $S = \{(x_1, x_2) = (0.4, 0.2)\}$. And the optimal value is $f = 0.6$. Below is the graph where we can see the optimal solution, i.e. the intersection of the blue line and the Feasible Set.

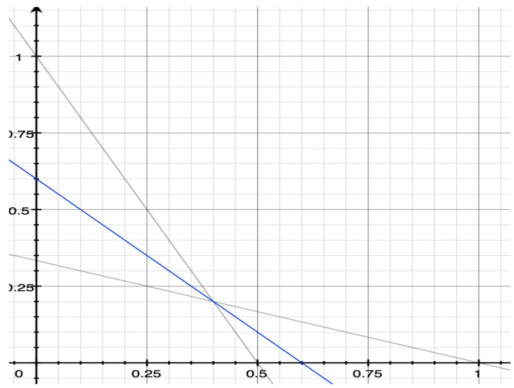

Figure 2: Set of Optimal Solutions for 2(a)

(b)

The set of optimal solutions is where either x_1 or x_2 goes to infinity, and the optimal value is just minus infinity. We cannot represent the optimal solutions with a graph but we know where it is.

(c)

Graphically, the set of optimal solutions is a set of points, i.e. $S = \{(x_1, x_2)|x_1 = 0, x_2 \geq 1\}$. And the optimal value is $f = 0$. Below is the graph where we can see the optimal solutions, i.e. the blue line above the point $(x_1, x_2) = (0, 1)$.

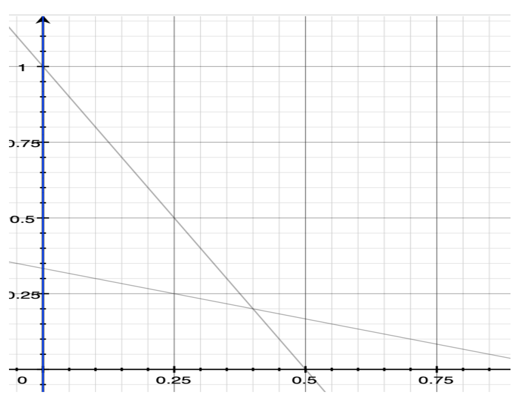

Figure 3: Set of Optimal Solutions for 2(c)

CVXPY Results for (a)-(e)

For the optimal variables, the first is x_1, the second is x_2.

```
Problem a
status: optimal
optimal value 0.5999999999116253
optimal var 0.3999999999724491 0.1999999999391762
Problem b
status: unbounded
optimal value -inf
optimal var None None
Problem c
status: optimal
optimal value -2.2491441767693296e-10
optimal var -2.2491441767693299e-10 1.5537158969947242
Problem d
status: optimal
optimal value 0.3333333330990559
optimal var 0.3333333334080862 0.3333333286259564
Problem e
status: optimal
optimal value 0.5000000000000003
optimal var 0.5000000000000001 0.1666666666666669
```

Figure 4: CVXPY Results for Problem 2

3

(a)

We introduce a new variable $\boldsymbol{t} \in \mathbb{R}^m$ and let $-\boldsymbol{t} \le \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \le \boldsymbol{t}$, where the inequality is defined elementwise. And note that for the constraint $||\boldsymbol{x}||_\infty \le 1$, it is equivalent to $-\mathbf{1} \le \boldsymbol{x} \le \mathbf{1}$(also elementwise). Then the original problem can be transformed into

$$\begin{aligned} &\min_{\boldsymbol{x},\boldsymbol{t}} \mathbf{1}^T \boldsymbol{t} \\ s.t. \quad & -\mathbf{1} \leq \boldsymbol{x} \leq \mathbf{1} \\ & -\boldsymbol{t} \leq \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \leq \boldsymbol{t} \end{aligned}$$

(b)

Note that the optimal variable is given in form of a row vector, whose transpose is the actual answer.

```
status: optimal
optimal value 13.999999988898507
optimal var [[ 1. -1.]]
```

Figure 5: CVXPY Results for Problem 3(b)

(c)

```
status: optimal
optimal value 13.999999998611605
optimal var x: [[ 1. -1.]] var t: [[4. 6. 4.]]
```

Figure 6: CVXPY Results for Problem 3(c)

4

(a)

Since $\boldsymbol{X}$ has full column rank, $\boldsymbol{w}^*$ is simply $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$, which is $(1.5, 2)^T$. And the optimal value is 4.

(b)(c)

```
Lasso
Parameter t = 1
status: optimal
optimal value 8.999999833595501
optimal var [[9.99962136e-01 3.78500362e-05]]

Parameter t = 10
status: optimal
optimal value 3.9999999726481903
optimal var [[1.49999883 1.99999744]]

Ridge
Parameter t = 1
status: optimal
optimal value 7.85748959131565
optimal var [[0.86266947 0.50576813]]

Parameter t = 100
status: optimal
optimal value 3.9999999752956037
optimal var [[1.50000023 2.0000014 ]]
```

Figure 7: CVXPY Results for Problem 4

Lasso with $t = 1$ has different solution, and the solution has zero component. But Lasso with $t = 10$ has the same solution with no zero component. Ridge with $t = 1$ has different solution with no zero component, while with $t = 100$ the solution is the same with no zero compenent.