Homework 5

Zhen

October 30, 2021

Problem 1

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

Proof:

$$\nabla f = \boldsymbol{x}-\boldsymbol{x}_0$$

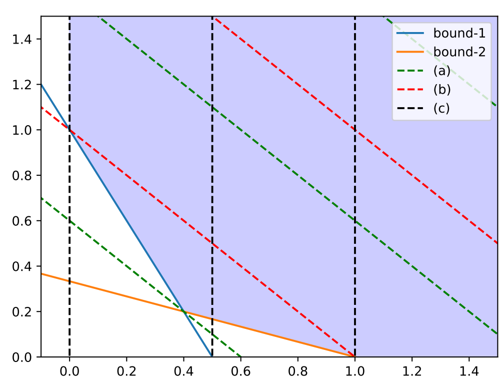
Also, the optimal point is unique because $\nabla^2 f > 0$.
Considering that: $\forall \boldsymbol{x}\in\bar{B}$,

$$\begin{aligned}
\nabla f(\frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|})^{\mathcal{T}}(\boldsymbol{x}-\frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}) &= (\frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}-\boldsymbol{x}_0)^{\mathcal{T}}(\boldsymbol{x}-\frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}) \\
&= (1-\frac{1}{\|\boldsymbol{x}_0\|})(\|\boldsymbol{x}_0\|-\boldsymbol{x}_0^{\mathcal{T}}\boldsymbol{x}) \\
&\geq (1-\frac{1}{\|\boldsymbol{x}_0\|})(\|\boldsymbol{x}_0\|-\|\boldsymbol{x}_0\|\times\|\boldsymbol{x}\|) \\
&\geq 0
\end{aligned}$$

By first-order optimality condition, we conclude that $\boldsymbol{x}^* = \frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}$.

Problem 2

The graph of question (a),(b) and (c) are as follows.



In the graph, the light blue area means the feasible set. And the series of dotted lines demonstrate the counter of function f.
So the solution for each questions is:
(a) $x_1 = 0.4, x_2 = 0.2, f = 0.6$
(a) $x_1 = -\infty, x_2 = -\infty, f = +\infty$, which means no optimal point.
(a) $x_1 = 0, x_2 = C(\geq 1), f = 0$

The code is at "P2.py".

```
------------(a)------------
status: optimal
optimal value: 0.5999999999116253
optimal var: x1 = 0.3999999999724491 , x2 = 0.1999999999391762
------------(b)------------
status: unbounded
optimal value: -inf
optimal var: x1 = None , x2 = None
------------(c)------------
status: optimal
optimal value: -2.2491441767693299e-10
optimal var: x1 = -2.2491441767693299e-10 , x2 = 1.5537158969947242
------------(d)------------
status: optimal
optimal value: 0.3333333334080862
optimal var: x1 = 0.3333333334080862 , x2 = 0.33333333286259564
------------(e)------------
status: optimal
optimal value: 0.5
optimal var: x1 = 0.5 , x2 = 0.16666666666666669
```

Problem 3

(a) $\boldsymbol{A} \overset{def}{=} (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)^{\mathcal{T}}, \boldsymbol{b} \overset{def}{=} (b_1, b_2, \dots, b_m)^{\mathcal{T}}, \boldsymbol{x} \overset{def}{=} (x_1, x_2, \dots, x_m)^{\mathcal{T}}$. Then we obtain

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_1 = \sum_{i=1}^{m} \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x} - b_i\right|$$

Also,

$$\begin{aligned} \|\boldsymbol{x}\|_\infty \le 1 &\iff \max_i |x_i| \le 1 \\ &\iff |x_i| \le 1,\ (\forall i \in \{1, \dots, n\}) \\ &\iff \boldsymbol{x} \le \mathbf{1} \wedge -\boldsymbol{x} \le \mathbf{1} \end{aligned}$$

Introducing $\boldsymbol{t} = (t_1, t_2, \dots, t_n)^{\mathcal{T}}$, we could reformulate the optimization problem as:

$$\begin{aligned} \min_{\boldsymbol{x}, \boldsymbol{t}} \quad & \mathbf{1}_m^{\mathcal{T}}\boldsymbol{t} \\ \textbf{s.t.} \quad & -\mathbf{1}_n \le \boldsymbol{x} \le \mathbf{1}_n \\ & \forall i \in \{1, 2, \dots, m\} : -t_i \le \boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x} - b_i \le t_i \end{aligned}$$

Rewrite the afformentioned optimal problem ($\boldsymbol{\omega} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{t} \end{bmatrix}$):

$$\min_{\boldsymbol{\omega}} \quad \begin{bmatrix} \mathbf{0}_n \\ \mathbf{1}_m \end{bmatrix}^{\mathcal{T}} \boldsymbol{\omega}$$

$$\textbf{s.t.} \quad \begin{bmatrix} \boldsymbol{E}_{n\times n} & \boldsymbol{O}_{n\times m} \\ -\boldsymbol{E}_{n\times n} & \boldsymbol{O}_{n\times m} \\ \boldsymbol{A} & -\boldsymbol{E}_{m\times m} \\ -\boldsymbol{A} & -\boldsymbol{E}_{m\times m} \end{bmatrix} \boldsymbol{\omega} \le \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_n \\ \boldsymbol{b} \\ -\boldsymbol{b} \end{bmatrix}$$

(b) The code is at "P3b.py".

```
status: optimal
optimal value: 13.999999990735517
optimal var: x1 = 1.0000000030380236 x2 = -1.0000000007971088
```

(c) The code is at “P3c.py”.

```
status: optimal
optimal value: 13.999999998611603
optimal var: x1 = 0.9999999999702551 x2 = -1.000000000224444
```

Problem 4

(a) $f(\boldsymbol{\omega}) \overset{def}{=} \|\boldsymbol{X}\boldsymbol{\omega} - \boldsymbol{y}\|_2^2$ and $\boldsymbol{\omega}^* = \arg\min_{\boldsymbol{\omega}} f(\boldsymbol{\omega})$. Considering that:

$$f(\boldsymbol{\omega}) = \boldsymbol{\omega}^{\mathcal{T}}\boldsymbol{X}^{\mathcal{T}}\boldsymbol{X}\boldsymbol{\omega} - 2\boldsymbol{y}^{\mathcal{T}}\boldsymbol{X}\boldsymbol{\omega} + \boldsymbol{y}^{\mathcal{T}}\boldsymbol{y}$$
$$\nabla f(\boldsymbol{\omega}) = 2\boldsymbol{X}^{\mathcal{T}}\boldsymbol{X}\boldsymbol{\omega} - 2\boldsymbol{X}^{\mathcal{T}}\boldsymbol{y}$$

By the fact that $\nabla f(\boldsymbol{\omega}^*) = 0$, we have

$$\boldsymbol{\omega}^* = \left(\boldsymbol{X}^{\mathcal{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathcal{T}}\boldsymbol{y} = \begin{bmatrix} 1.5 \\ 2 \end{bmatrix}$$

(b) The code is at “P4b.py”.

```
---------- 1 ----------
status: optimal
optimal value: 9.0000000633334
optimal var: x1 = 0.999962135922334 x2 = 3.785003616073247e-05
---------- 10 ----------
status: optimal
optimal value: 4.000000000012022
optimal var: x1 = 1.4999988302124445 x2 = 1.9999974407244057
```

When $t = 1$, the solution is **NOT** the same as that of (a) and has a zero component.
When $t = 10$, the solution is the same as that of (a) and does not contain a zero component.

(c) The code is at “P4c.py”.

```
---------- 1 ----------
status: optimal
optimal value: 7.857489685034338
optimal var: x1 = 0.8627047883360971 x2 = 0.5057078807022429
---------- 100 ----------
status: optimal
optimal value: 4.0000000000000195
optimal var: x1 = 1.500000021429177 x2 = 2.0000001348138836
```

When $t = 1$, the solution is **NOT** the same as that of (a) and does not has a zero component.
When $t = 10$, the solution is the same as that of (a) and does not contain a zero component.