11.1 Homework 5

1. $P_C(x) = \hat{x}$, $\|x - \hat{x}\| = dist(x, C) = \inf_{z \in C} \|x - z\|$

The finding of a projection is the result of the following problem

$\min_{x \in \bar{B}} f(x) = \frac{1}{2}\|x - x_0\|^2$

here $x_0 \notin \bar{B}$, according to the first-order optimal condition:

$\nabla f(x^*)^T \cdot (x - x^*) \geq 0$, that is : $(x^* - x_0)^T (x - x^*) \geq 0$

let $x^* = \frac{x_0}{\|x_0\|}$, $(x^* - x_0)^T (x - x^*) = (\frac{1}{\|x_0\|} - 1)\, x_0^T x - (\frac{1}{\|x_0\|} - 1) \cdot \frac{1}{\|x_0\|} x_0^T x_0$

$= (\frac{1}{\|x_0\|} - 1)\, x_0^T (x - \frac{x_0}{\|x_0\|})$

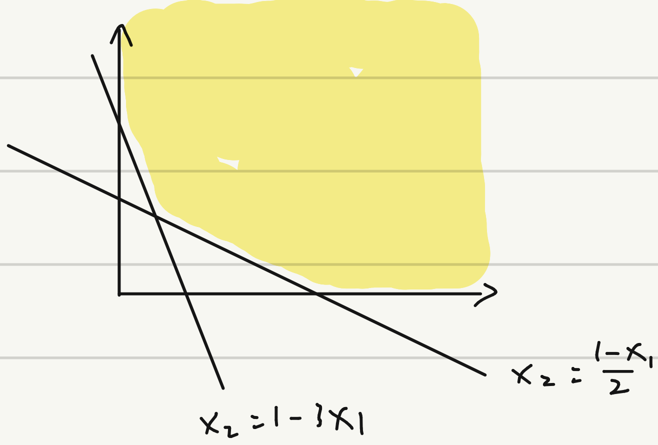
$\frac{x_0}{\|x_0\|}$ is also the gradient on $\partial \bar{s}$ at $\frac{x_0}{\|x_0\|}$, so $\frac{x}{\|x_0\|}$ is also a hyperplane that

$w = \frac{x_0}{\|x_0\|}$, $(x - \frac{x_0}{\|x_0\|})^T \frac{x_0}{\|x_0\|} \leq 0$,

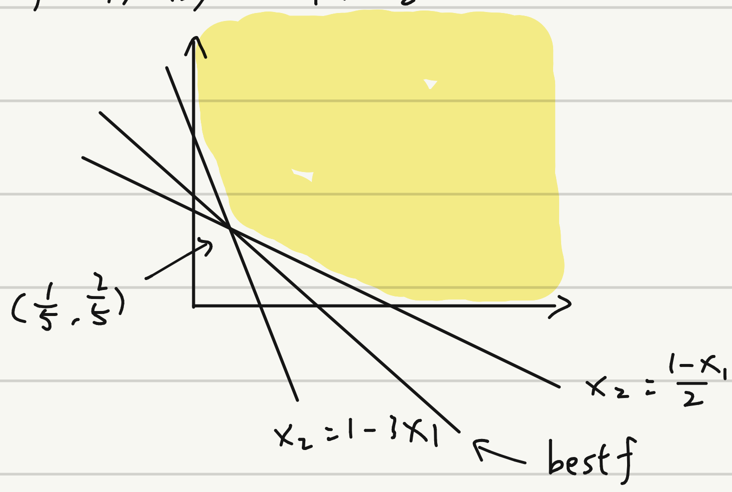
$\|x_0\| > 1$, $\frac{1}{\|x_0\|} < 0$, so $(x^* - x_0)^T (x - x^*) \geq 0$

the $P_{\bar{s}}(x_0)$ is unique. $P_{\bar{s}}(x_0)$ is exactly $\frac{x_0}{\|x_0\|}$

2. the feasible set looks like:



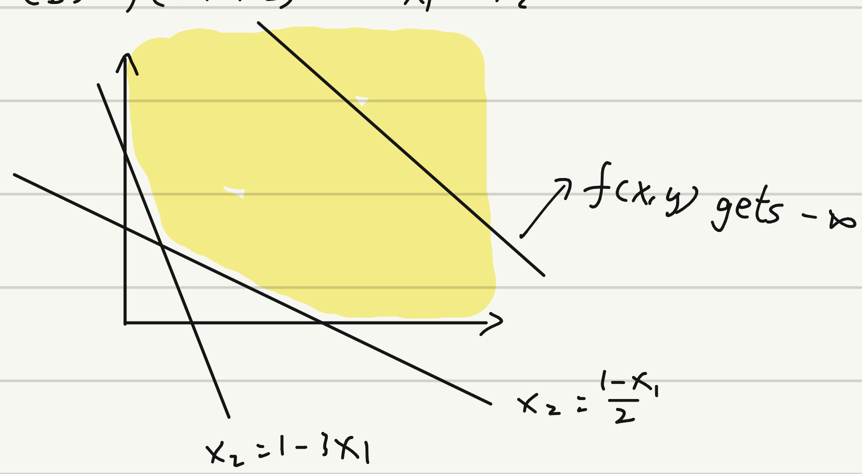
(a) $f(x_1, x_2) = x_1 + x_2$



$f(x_1, x_2)_{min} = f(\frac{1}{5}, \frac{2}{5}) = \frac{3}{5}$

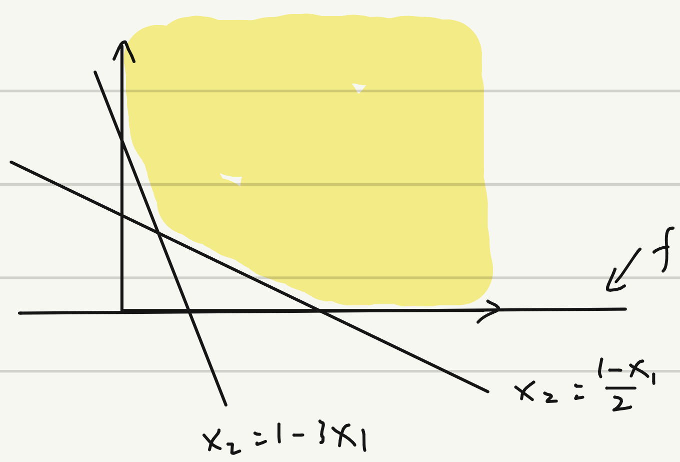
solution: $\{(\frac{1}{5}, \frac{2}{5})\}$

(b) $f(x_1, x_2) = -x_1 - x_2$



solution : $\emptyset$

(c) $f(x_1, x_2) = x_1$



we have $x_1 \geq 0$, and so

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

the results of cvxpy for (a)(b)(c) matches my answer

```
Problem(a): status: optimal
optimal value 0.5999999999116254
optimal var: x1 = 0.1999999999391762 x2 = 0.3999999999724492
Problem(b): status: unbounded
optimal value -inf
optimal var: x1 = None x2 = None
Problem(c): status: optimal
optimal value -1.232214801046685e-10
optimal var: x1 = -1.232214801046685e-10 x2 = 1.7673174212389093
Problem(d): status: optimal
optimal value 0.3333333334080862
optimal var: x1 = 0.33333333286259564 x2 = 0.333333333334080862
Problem(e): status: optimal
optimal value 0.6923076923076925
optimal var: x1 = 0.6923076923076924 x2 = 0.15384615384615388
```

3. (a) $\|Ax - b\| = \max (Ax - b)_i$ s.t. $x_i \leq 1$

$= \max_{1 \leq i \leq m} (a_i^T x - b_i$

can be transfomulate into LP as:

$\min_{x, t} t$ s.t. $t \geq |a_i^T x - b_i|, |x_i| \leq 1$, $1 \leq i \leq m$

(b) Result of cvxpy is

Problem(b): status: optimal
optimal value 5.333333333553781
optimal var: x = [[-0.33333333] [0.33333333]]
optimal

(c) Result is

Problem(c): status: optimal
optimal value 5.333333333260568
optimal var: x = [[-0.33333333] [0.33333333]]
optimal

the same

4. (a) Solution: the normal equation is $X^T X w = X^T y$

$$w = (X^T X)^{-1} X^T y \approx \begin{pmatrix} 1.22 \\ -0.21 \\ 0.16 \\ -0.46 \\ 1.19 \\ 0.006 \end{pmatrix}$$

(b) It's not the same, for $\|w\|_1 > 1$ in (a). When $t=10$ it's the same

the solution is not sparce when $X^T X$ is full-rank,

but may be sparce if $X^T X$ is not full-rank

(c) not the same, for the same reason when $\|w\|_2 = 1$, and same when $\|w\|_2$

no zero components, it seems

Problem(b) when t = 1: status: optimal
optimal value 31.314550054478023
optimal var: w = [[5.54241960e-01]
[4.31525539e-09]
[9.92071629e-10]
[9.38255329e-09]
[4.30602870e-01]
[1.51551568e-02]]
optimal
Problem(b) when t = 10: status: optimal
optimal value 13.295569218508422
optimal var: w = [[1.22171615]
[-0.21469843]
[0.15549443]
[-0.45868521]
[1.18537859]
[0.00613412]]
optimal
Problem(c) when t = 1: status: optimal
optimal value 16.17313072100805
optimal var: w = [[0.52519352]
[0.08615554]
[0.09403397]
[0.12511932]
[0.82964038]
[0.06283835]]
optimal
Problem(c) when t = 100:
status: optimal
optimal value
13.295569218196661
optimal var: w = [[1.22170661]
[-0.21469307]
[0.15549204]
[-0.4586777]
[1.18537706]
[0.00613317]]
optimal