CS257 Linear and Convex Optimization Homework 10

Xiangyuan Xue (521030910387)

1. (a) Suppose $\boldsymbol{x}$ contains n components, denoted by $x_1, x_2, \dots, x_n$. The approximating equality constrained problem is

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & F(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} - \frac{1}{t}\sum_{i=1}^{n} \log x_i \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

(b) The gradient of $F(\boldsymbol{x})$ is

$$\nabla F(\boldsymbol{x}) = \boldsymbol{c} - \frac{1}{t}\begin{pmatrix} \frac{1}{x_1} \\ \frac{1}{x_2} \\ \vdots \\ \frac{1}{x_n} \end{pmatrix}$$

The Hessian matrix of $F(\boldsymbol{x})$ is

$$\nabla^2 F(\boldsymbol{x}) = \frac{1}{t}\begin{pmatrix} \frac{1}{x_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n^2} \end{pmatrix}$$

(c) The function `centering_step` is as follow

```
def centering_step(c, A, b, x0, t):
    def F(x): return (c @ x - np.sum(np.log(np.maximum(x, 0))) / t
    )
    def Fp(x): return (c - (1 / x) / t)
    def Fpp(x): return (np.diag(1 / (x ** 2)) / t)
    x_traces = nt.newton_eq(F, Fp, Fpp, x0, A, b)
    return x_traces[-1]
```

The function `barrier` is as follow

```
def barrier(c, A, b, x0, tol=1e-8, t0=1, rho=10):
    t = t0
    x = np.array(x0)
```

```python
    x_traces = [np.array(x0)]
    lim = len(b) / tol
    while t < lim:
        x = centering_step(c, A, b, x, t)
        x_traces.append(np.array(x))
        t = rho * t
    return x_traces
```

(d) The standard form is written as

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & -3x_1 - x_2 \\
\text{s.t.} \quad & x_1 + 2x_2 + x_3 = 8 \\
& x_1 - x_2 + x_4 = 3 \\
& x_1, x_2, x_3, x_4 \geq 0
\end{aligned}
$$

where the initial point is set to be $\boldsymbol{x}_0 = (2, 1, 4, 2)^T$. The output is as follow:

```
iteration 0: [2 1 4 2]
iteration 1: [4.05162269 1.66475112 0.61887506 0.61312843]
iteration 2: [4.60203415 1.66223524 0.07349538 0.06020108]
iteration 3: [4.66017038 1.66617243 0.00748475 0.00600205]
iteration 4: [4.66601670e+00 1.66661672e+00 7.49847335e-04
   6.00020561e-04]
iteration 5: [4.66660167e+00 1.66666167e+00 7.49984725e-05
   6.00002051e-05]
iteration 6: [4.66666017e+00 1.66666617e+00 7.49923073e-06
   5.99939942e-06]
iteration 7: [4.66666602e+00 1.66666662e+00 7.49924165e-07
   5.99939480e-07]
iteration 8: [4.66666660e+00 1.66666666e+00 7.42466142e-08
   5.93972957e-08]
iteration 9: [4.66666666e+00 1.66666667e+00 4.99833499e-09
   3.99866871e-09]
optimal value: -15.666666653337767
```

The projected trajectory is plotted as follow

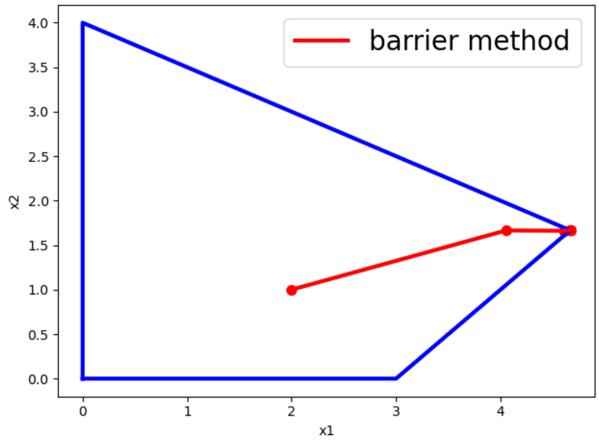

Figure 1: trajectory for barrier method

Therefore, the optimal solution is $\boldsymbol{x}^* = (\frac{14}{3}, \frac{5}{3}, 0, 0)^T$ and the optimal value is $f^* = -\frac{47}{3}$.

2. (a) The dual LP in the standard form is written as

$$\begin{aligned} \max_{\boldsymbol{\mu}} \quad & -8\mu_1 - 3\mu_2 \\ \text{s.t.} \quad & -\mu_1 - \mu_2 + \mu_3 = -3 \\ & -2\mu_1 + \mu_2 + \mu_4 = -1 \\ & \mu_1, \mu_2, \mu_3, \mu_4 \geq 0 \end{aligned}$$

(b) The symmetric dual LP is written as

$$\begin{aligned} \max_{\boldsymbol{\mu}} \quad & -8\mu_1 - 3\mu_2 \\ \text{s.t.} \quad & -\mu_1 - \mu_2 \leq -3 \\ & -2\mu_1 + \mu_2 \leq -1 \\ & \mu_1, \mu_2 \geq 0 \end{aligned}$$

(c) Graphically, the dual LP is plotted as follow

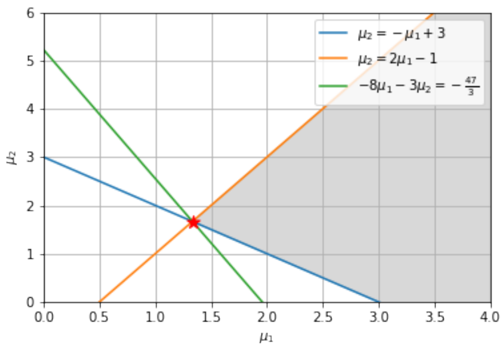

Figure 2: visualization of dual LP

Therefore, the dual optimal solution is $\boldsymbol{\mu}^* = (\frac{4}{3}, \frac{5}{3})$ and the dual optimal value is $f^* = -\frac{47}{3}$, which is the same as the primal optimal value.

(d) The standard form is written as

$$\begin{aligned}
\min_{\boldsymbol{\mu}} \quad & 8\mu_1 + 3\mu_2 \\
\text{s.t.} \quad & -\mu_1 - \mu_2 + \mu_3 = -3 \\
& -2\mu_1 + \mu_2 + \mu_4 = -1 \\
& \mu_1, \mu_2, \mu_3, \mu_4 \geq 0
\end{aligned}$$

where the initial point is set to be $\boldsymbol{\mu}_0 = (4, 1, 2, 6)^T$. The output is as follow:

```
iteration 0: [4. 1. 2. 6.]
iteration 1: [1.61583503 1.63097966 0.24681469 0.60069039]
iteration 2: [1.36062983 1.6610997  0.02172952 0.06015996]
iteration 3: [1.33604921 1.66609664 0.00214584 0.00600178]
iteration 4: [1.33360478e+00 1.66660954e+00 2.14315562e-04
   6.00017974e-04]
iteration 5: [1.33336047e+00 1.66666095e+00 2.14267293e-05
   5.99941600e-05]
iteration 6: [1.33333605e+00 1.66666610e+00 2.14264404e-06
   5.99939647e-06]
iteration 7: [1.33333360e+00 1.66666661e+00 2.12133289e-07
   5.93973007e-07]
iteration 8: [1.33333336e+00 1.66666666e+00 1.90444581e-08
   5.33244605e-08]
iteration 9: [1.33333333e+00 1.66666667e+00 2.68017940e-10
   7.50474575e-10]
dual optimal value: -15.666666669168201
```

Therefore, the dual optimal solution is $\boldsymbol{\mu}^* = (\frac{4}{3}, \frac{5}{3}, 0, 0)^T$ and the dual optimal value is $f^* = -\frac{47}{3}$.

3. (a) Since f is monotonically increasing on $\mathbb{R}$, the minimum is reached when $x = 0$. Therefore, the optimal solution is $x^* = 0$ and the optimal value is $f^* = \log 3$.

(b) The Lagrangian is

$$\mathcal{L}(x, \mu) = \log(2 + e^x) - \mu x$$

The dual function is

$$\phi(\mu) = \inf_{x \in \mathbb{R}} \mathcal{L}(x, \mu) = \inf_{x \in \mathbb{R}} \left[\log(2 + e^x) - \mu x\right]$$

Let the gradient be zero

$$\frac{\partial}{\partial x} \mathcal{L}(x, \mu) = \frac{e^x}{2 + e^x} - \mu = 0$$

which yields

$$x = \log \frac{2\mu}{1-\mu}$$

Note that if $\mu < 0$, $\mathcal{L}$ is monotonically increasing and $\phi(\mu) = \lim_{x\to-\infty} \mathcal{L}(x,\mu) = -\infty$. If $\mu > 1$, $\mathcal{L}$ is monotonically decreasing and $\phi(\mu) = \lim_{x\to+\infty} \mathcal{L}(x,\mu) = -\infty$. If $\mu = 0$, we have $\phi(\mu) = \log 2$. If $\mu = 1$, we have $\phi(\mu) = 0$. Otherwise, $\phi(\mu) = \mathcal{L}\left(\log \frac{2\mu}{1-\mu}, \mu\right) = -\mu \log \mu - (1-\mu)\log \frac{1-\mu}{2}$. Thus, the dual function is

$$\phi(\mu) = \begin{cases} -\mu \log \mu - (1-\mu)\log \dfrac{1-\mu}{2}, & 0 \le \mu \le 1 \\ -\infty, & \text{otherwise} \end{cases}$$

where we suppose $0 \log 0 = 0$ for convenience. The dual problem is

$$\begin{aligned} \max_{\mu} \quad & \phi(\mu) = \begin{cases} -\mu \log \mu - (1-\mu)\log \dfrac{1-\mu}{2}, & 0 \le \mu \le 1 \\ -\infty, & \text{otherwise} \end{cases} \\ \text{s.t.} \quad & \mu \ge 0 \end{aligned}$$

(c) We might as well let $\mu \in (0,1)$. Let the gradient be zero

$$\frac{\partial}{\partial \mu}\phi(\mu) = \log \frac{1-\mu}{2\mu} = 0$$

which yields $\mu = \frac{1}{3}$ and $\phi(\frac{1}{3}) = \log 3$. Thus

$$\phi^* = \max\left\{\phi(0), \phi(1), \phi\left(\frac{1}{3}\right)\right\} = \log 3$$

Therefore, the dual optimal solution is $\mu^* = \frac{1}{3}$ and the dual optimal value is $\phi^* = \log 3$, which is the same as the primal optimal value, indicating that strong duality holds.

4. (a) The Lagrangian is

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu}) &= x_1^2 + x_2^2 + \mu_1\left[(x_1-2)^2 + (x_2-1)^2 - 1\right] \\ &\quad + \mu_2\left[(x_1-2)^2 + (x_2+1)^2 - 1\right] \\ &= (1+\mu_1+\mu_2)x_1^2 - 4(\mu_1+\mu_2)x_1 \\ &\quad + (1+\mu_1+\mu_2)x_2^2 - 2(\mu_1-\mu_2)x_2 + 4(\mu_1+\mu_2) \end{aligned}$$

If $\mu_1 + \mu_2 > -1$, $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu})$ reaches minimum when

$$\begin{cases} x_1 = \dfrac{2(\mu_1+\mu_2)}{1+\mu_1+\mu_2} \\ x_2 = \dfrac{\mu_1-\mu_2}{1+\mu_1+\mu_2} \end{cases}$$

which yields

$$\inf_{\boldsymbol{x}\in\mathbb{R}^2} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu}) = \frac{4(\mu_1+\mu_2) - (\mu_1-\mu_2)^2}{1+\mu_1+\mu_2}$$

Otherwise, $\mathcal{L}(\boldsymbol{x},\boldsymbol{\mu})$ is unlimited and $\inf_{\boldsymbol{x}\in\mathbb{R}^2}\mathcal{L}(\boldsymbol{x},\boldsymbol{\mu})=-\infty$. Thus, the dual function is

$$\phi(\boldsymbol{\mu})=\begin{cases}\dfrac{4(\mu_1+\mu_2)-(\mu_1-\mu_2)^2}{1+\mu_1+\mu_2}, & \mu_1+\mu_2>-1\\ -\infty, & \mu_1+\mu_2\le -1\end{cases}$$

Since it holds that $\boldsymbol{\mu}\ge\mathbf{0}$, the dual problem is

$$\begin{aligned}\max_{\boldsymbol{\mu}}\quad & \varphi(\boldsymbol{\mu})=\frac{4(\mu_1+\mu_2)-(\mu_1-\mu_2)^2}{1+\mu_1+\mu_2}\\ \text{s.t.}\quad & \boldsymbol{\mu}\ge\mathbf{0}\end{aligned}$$

(b) Note that $\boldsymbol{x}=(2,0)$ is the only feasible point for the primal problem, so $f^*=f(2,0)=4$. By weak duality, we know $\phi^*\le f^*=4$. On the other hand

$$\lim_{\substack{\mu_1=\mu_2\\ \mu_1\to+\infty}}\phi(\mu_1,\mu_2)=\lim_{\mu'\to+\infty}\frac{4\mu'}{1+\mu'}=4$$

Therefore

$$\phi^*=\sup_{\boldsymbol{\mu}\ge\mathbf{0}}\phi(\boldsymbol{\mu})=4=f^*$$

which indicates that strong duality holds.

(c) Since $(2,0)$ is the only feasible point for the primal problem, an interior point does not exist and the problem is not strictly feasible. In other words, Slater's condition does not hold.

However, strong duality holds for this problem. We can conclude that Slater's condition is not necessary for strong duality.

(d) Since $\boldsymbol{\mu}\ge\mathbf{0}$, we have

$$\phi(\boldsymbol{\mu})=\frac{4(\mu_1+\mu_2)-(\mu_1-\mu_2)^2}{1+\mu_1+\mu_2}\le\frac{4(\mu_1+\mu_2)}{1+\mu_1+\mu_2}<4$$

Thus, the dual optimal value $\phi^*=4$ is not attained by any dual feasible point, which indicates that the KKT conditions do not hold.

Consider part of the KKT conditions

$$\begin{cases}\dfrac{\partial\mathcal{L}}{\partial x_1}=2(1+\mu_1+\mu_2)x_1-4(\mu_1+\mu_2)=0\\ \dfrac{\partial\mathcal{L}}{\partial x_2}=2(1+\mu_1+\mu_2)x_2-2(\mu_1-\mu_2)=0\end{cases}$$

At $\boldsymbol{x}^*=(2,0)$, no corresponding Lagrange multipliers exist. We can also find the gradients of the constraints at $\boldsymbol{x}^*$ are $(0,-2)^T$ and $(0,2)^T$, indicating that $\boldsymbol{x}^*$ is not a regular point, which verifies that the KKT conditions do not hold.

5. (a) The dual function is

$$\begin{aligned}\phi(\mu)&=\inf_{\boldsymbol{x}\ge\mathbf{0}}\left\{x_1^5+x_2^5+\mu(1-x_1-x_2)\right\}\\ &=\inf_{x_1\ge0}\left\{x_1^5-\mu x_1\right\}+\inf_{x_2\ge0}\left\{x_2^5-\mu x_2\right\}+\mu\end{aligned}$$

Consider the function $\lambda(x) = x^5 - \mu x$ where $x \geq 0$. If $\mu < 0$, the minimum is reached when $x = 0$, then $\lambda(0) = 0$. If $\mu \geq 0$, let the gradient $\lambda'(x) = 5x^4 - \mu = 0$ and we have $x = \sqrt[4]{\frac{\mu}{5}}$, then $\lambda\left(\sqrt[4]{\frac{\mu}{5}}\right) = -4\left(\frac{\mu}{5}\right)^{\frac{5}{4}}$. So the explicit expression is

$$\phi(\mu) = \begin{cases} -8\left(\dfrac{\mu}{5}\right)^{\frac{5}{4}} + \mu, & \mu \geq 0 \\ \mu, & \mu < 0 \end{cases}$$

(b) Consider $\mu \geq 0$. Let the gradient be zero

$$\phi'(\mu) = -2\left(\frac{\mu}{5}\right)^{\frac{1}{4}} + 1 = 0$$

which yields $\mu = \frac{5}{16}$ and a maximal $\phi\left(\frac{5}{16}\right) = \frac{1}{16}$. Thus

$$\phi^* = \max\left\{\phi(0), \phi\left(\frac{5}{16}\right)\right\} = \frac{1}{16}$$

Therefore, the dual optimal solution is $\mu^* = \frac{5}{16}$ and the dual optimal value is $\phi^* = \frac{1}{16}$.

(c) For $\boldsymbol{x} = (1, 1) \in D$, we have $1 - x_1 - x_2 < 0$, so the primal problem is convex and strictly feasible, so Slater's condition holds. By strong duality, we know that the primal optimal value $f^* = \phi^* = \frac{1}{16}$.

(d) The dual function is

$$\phi_1(\boldsymbol{\mu}) = \inf_{\boldsymbol{x} \in \mathbb{R}^2} \left\{x_1^5 + x_2^5 - \mu_1(x_1 + x_2 - 1) - \mu_2 x_1 - \mu_3 x_2\right\} = -\infty$$

The dual optimal value is $\phi_1^* = -\infty$, which is different from $f_1^* = f^* = \frac{1}{16}$. Therefore, the strong duality does not hold for this equivalent problem.