CSCI 5254 Homework 4

Tuguluke Abulitibu

October 21, 2020

Chapter 4, Basic definitions

5.1

(a)

- feasible set: $[2, 4]$
- optimal value: $x^* = 2$
- optimal solution: $p^* = 2^2 + 1 = 5$

(b)

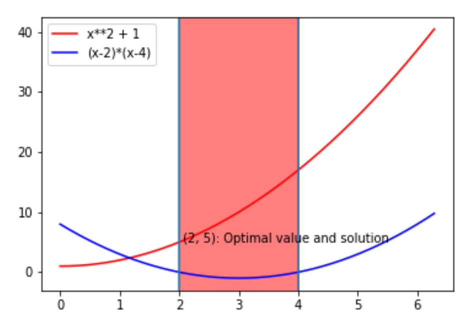


Figure 1: feasible set and optimal

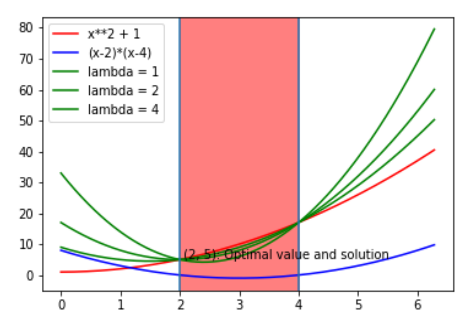


Figure 2: with Lagrange

Lagrangian:

$$L(x, \lambda) = x^2 + 1 - \lambda(x-2)(x-4) = (1+\lambda)x^2 - 6\lambda x + 1 + 8\lambda$$

let $(-\dfrac{9\lambda^2}{1+\lambda} + 1 + 8\lambda)' = 0$. we get $\lambda = 2$. also from Figure 2, we can easily see that lower bound property $5 = p^* \geq \inf_x L(x, \lambda)$ holds, in fact from Figure 3 we can see it hold at $\lambda = 2$

with $L'(x, \lambda) = 2(1+\lambda)x - 6\lambda = 0$, we have $x = \dfrac{3\lambda}{1+\lambda}$ Plug back in, we get dual function

$$= (1+\lambda)(\frac{3\lambda}{1+\lambda})^2 - 6\lambda\frac{3\lambda}{1+\lambda} + 1 + 8\lambda = \frac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda$$

hence

$$g(\lambda) = \begin{cases} \dfrac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda & 1+\lambda > 0 \\ -\infty & 1+\lambda \leq 0 \end{cases}$$

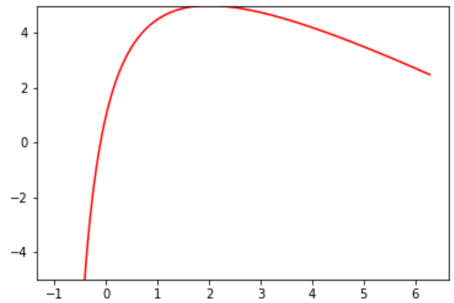


Figure 3: λ plot zoom in

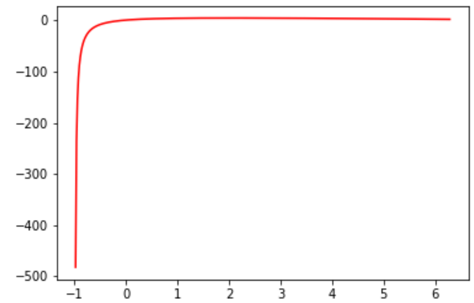


Figure 4: λ plot

(c)

By textbook definition (5.16)

$$\boxed{\begin{array}{ll} \text{maximize} & \dfrac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda \\ \text{subject to} & \lambda \succeq 0 \end{array}}$$

since $\nabla(\dfrac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda) = -\dfrac{18}{(x+1)^3}$, hence it is concave.
From part (b), we already know $\lambda^* = 2$, and the strong duality holds.

(d)

Figure 5: $p^*(u)$

From Figure 1 we can see that $\min f(x) = (x-2)(x-4) = -1$, hence $p^*(u) = -\infty$ if $u < -1$.
From $x^2 - 6x + 8 - u = 0$, with quadratic formula, we got the solution: $x = \dfrac{6 \pm \sqrt{4+4u}}{2}$. Hence $[3-\sqrt{1-u}, 3+\sqrt{1-u}]$, we can see $x^8 = 3 - \sqrt{1-u}$

$$p^*(u) = \begin{cases} -\infty & u < -1 \\ x^2 + 1 = (3-\sqrt{1-u})^2 + 1 & -1 \le u \le 8 \\ \inf(x^2+1) = 1 & u \ge 8 \end{cases}$$

Finally:

$$\frac{dp}{du} = 1 - \frac{3}{\sqrt{1+u}} = -2(\text{ when } u = 0) = -\lambda^*$$

Chapter 4, Examples and applications

5.11

With hint $y_i = A_i^T x + b_i$, the problem transforms into

$$\boxed{\begin{array}{ll} \text{min} & \sum_i^N \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 \\ \text{subject to} & A_i x + b = y_i \end{array}}$$

hence

$$L(x,y,\nu) = \sum_i^N \|y_i\|_2 + \frac{1}{2}\|x-x_0\|_2^2 + \sum_i^N \nu_i^T(A_i x + b_i - y_i)$$

$$= \underbrace{\sum_i^N \|y_i\|_2}_{\text{Conjugate of a norm}} + \underbrace{\frac{1}{2}\|x-x_0\|_2^2 + (\sum_i \nu_i A_i)x}_{\text{can be min by derivative}} + b^T\nu$$

since $\sum_i^N \|y_i\|_2 \begin{cases} 0, & \|y_i\|_2 \le 1 \\ -\infty, & \|y_i\|_2 > 1 \end{cases}$[1], also $[\frac{1}{2}\|x-x_0\|_2^2 + (\sum_i \nu_i A_i)x]' = 0 \Rightarrow x = x_0 - \sum_i \nu_i A_i$, plug both back, we have the dual function

$$g(y_i) = \begin{cases} \sum_i^N \|A_i^T x_0 + b_i\|_2 - \frac{1}{2}\|\nu_i A_i\|_2^2 & \|y\| \le 1 \\ -\infty, & \mathit{otherwise} \end{cases}$$

hence the dual problem:

$$\boxed{\begin{array}{ll} \text{maximize} & \sum_i^N \|A_i^T x_0 + b_i\|_2 - \frac{1}{2}\|A_i \nu_i\|_2^2 \\ \text{subject to} & \|y_i\|_2 \le 1,\ i = 1, 2, \dots, N \end{array}}$$

5.13

(a)

$$L(x,\mu,\nu) = c^T x + \lambda(Ax - b) - \sum_i \nu_i(x_i - x_i^2)$$

$$= c^T x + \lambda(Ax-b) - \nu^T x + x^T \mathbf{diag}(\nu)x$$

$$= x^T\mathbf{diag}(\nu)x + (C + A^T\lambda - \nu)^T x - b^T\lambda$$

by letting $L' = 2x\mathbf{diag}(\nu) + [C + A^T\lambda - \nu]^T = 0$, we have $x = -\dfrac{(C + A^T\lambda - \nu)^T}{2\mathbf{diag}(\nu)}$, plug back in, we have

$$g(x) = \frac{(C + A^T\lambda - \nu)^T}{2diag(\nu)}\mathbf{diag}(\nu)\frac{(C + A^T\lambda - \nu)}{2\mathbf{diag}(\nu)} - (C + A^T\lambda - \nu)\frac{(C + A^T\lambda - \nu)}{2\mathbf{diag}(\nu)} - b^T\lambda$$

$$= -b^T\lambda - \frac{1}{4\mathbf{diag}(\nu)}(C + A^T\mu - \nu)^2$$

since $\sum_i \nu_i x_i^2 = X^t \begin{bmatrix} \nu_1 & & \\ & \ddots & \\ & & \nu_n \end{bmatrix} X$ is bounded only when $\nu_i \ge 0, \forall i$, then we get the dual function

$$g(\lambda,\nu) = \begin{cases} -b^T\lambda - \frac{1}{4}\sum_i \dfrac{(c_i + a_i^T\lambda_i - \nu_i)^2}{\nu_i} & \nu \ge 0 \\ -\infty & \text{Otherwise} \end{cases}$$

[1]Example 3.26 on page 93

Dual problem

$$\boxed{\begin{array}{ll} \text{maximize} & -b^T\lambda - \frac{1}{4}\sum_i \dfrac{(c_i + a_i^T\lambda_i - \nu_i)^2}{\nu_i} \\ \text{subject to} & \nu \succeq 0 \end{array}}$$

Since $\sup_{\nu_i} -\dfrac{(c_i + a_i^T\lambda_i - \nu_i)^2}{\nu_i} \begin{cases} c_i + a_i^T\lambda_i & c_i + a_i^T\lambda_i \leq 0 \\ 0 & c_i + a_i^T\lambda_i > 0 \end{cases}$, therefore we can rewrite the problem as:

$$\boxed{\begin{array}{ll} \text{maximize} & -b^T\lambda + \sum_i \min\{0, (c_i + a_i^T\lambda_i)\} \\ \text{subject to} & \lambda \succeq 0 \end{array}}$$

Which will come in handy in part (b)

(b)

Let us copy the LP-relaxation, and since

$$0 \leq x_i \leq 1 \leftrightarrow \begin{cases} -x_i \leq 0 \\ x_i - 1 \leq 0 \end{cases}$$

therefore[2]

$$L = c^T + \mu^T(Ax - b) - \nu^T x + w^T(x - \mathbf{1})$$

which is linear (bounded when coefficient of x is zero), hence we have

$$g(\mu, \nu, w) = \begin{cases} -\mu^T b - w^T\mathbf{1}; & c + A^T\mu - \nu + w = 0 \\ -\infty; & otherwise \end{cases}$$

convert it into dual problem:

$$\boxed{\begin{array}{ll} \text{maximize} & -b^T\mu - \mathbf{1}^T w \\ \text{subject to} & c + A^T\mu - \nu + w = 0 \\ & \mu, \nu, w \succeq 0 \end{array}}$$

let $\nu = 0$:

$$\boxed{\begin{array}{ll} \text{maximize}_{\mu,w} & -b^T\mu - \mathbf{1}^T w \\ \text{subject to} & c + A^T\mu + w = 0 \\ & \mu, w \succeq 0 \end{array}}$$

further, we make it into solve for w:

$$\boxed{\begin{array}{ll} \text{maximize}_{w} & -b^T\mu - \mathbf{1}^T w \\ \text{subject to} & c + A^T\mu + w = 0 \\ & w \succeq 0 \end{array}}$$

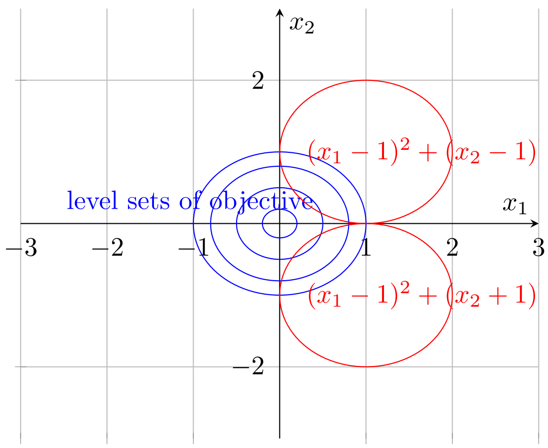
Here we can see this is the same as part(a), hence **the lower bound obtained via lagrangian relaxation and via the LP relaxation are the same**.

[2] $\mathbf{1} = [1, 1, \ldots, 1]^T$

Chapter 4, Optimality conditions

5.26

(a)



$x^* = (1, 0)$, and $p^* = 0^2 + 1^2 = 1$

(b)

- primal feasibility: $\begin{cases}(x_1-1)^2+(x_2-1)^2 \le 1\\ (x_1-1)^2+(x_2+1)^2 \le 1\end{cases}$
- dual feasibility : $\begin{cases}\lambda_1 \ge 0\\ \lambda_2 \ge 0\end{cases}$
- complimentary slackness: $\begin{cases}\lambda_1[(x_1-1)^2+(x_2-1)^2-1]=0\\ \lambda_2[(x_1-1)^2+(x_2+1)^2-1]=0\end{cases}$
- first order condition: $\begin{cases}2x_1+2(x_1-1)\lambda_1+2(x_1-1)\lambda_2=0\\ 2x_2+2(x_2-1)\lambda_1+2(x_2+1)\lambda_2=0\end{cases}$

plug $x^* = (1, 0)$ into KKT, we have

$$\begin{cases}1 \le 1\\ \lambda_{1,2} \ge 0\\ 2=0\\ 0=0\\ -2\lambda_1+2\lambda_2=0\end{cases}$$

the dual optimal is not obtained, which just show that there **does not exist λ_1^*, λ_2^* that x^* is optimal.**

(c)

$$L(x_1,x_2,\lambda_1,\lambda_2) = x_1^2+x_2^2+\lambda_1[(x_1-1)^2+(x_2-1)^2-1]+\lambda_2[(x_1-1)^2+(x_2+1)^2-1]$$

$$\begin{aligned}L(x_1,x_2,\lambda_1,\lambda_2) &= x_1^2+x_2^2+\lambda_1[(x_1^2-2x_1+1)+(x_2^2-2x_2+1)-1]+\lambda_2[(x_1^2-x_1+1)+(x_2^2+2x_2+1)-1]\\ &= (1+\lambda_1+\lambda_2)x_1-2(\lambda_1+\lambda_2)x_1+(1+\lambda_1+\lambda_2)x_2-2(\lambda_1-\lambda_2)x_2+\lambda_1+\lambda_2\end{aligned}$$

by letting

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 2x_1(1+\lambda_1+\lambda_2) - 2(\lambda_1+\lambda_2) = 0 \\ \dfrac{\partial L}{\partial x_1} = 2x_2(1+\lambda_1+\lambda_2) - 2(\lambda_1-\lambda_2) = 0 \end{cases}$$

we get $x^* = (\dfrac{\lambda_1+\lambda_2}{1+\lambda_1+\lambda_2}, \dfrac{\lambda_1-\lambda_2}{1+\lambda_1+\lambda_2})$, plug back in, we will get dual function

$$g(\lambda_1,\lambda_2) = \begin{cases} \dfrac{-2(\lambda_1-\lambda_2)^2 - 2(\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + \lambda_1+\lambda_2 & 1+\lambda_1+\lambda_2 \geq 0 \\ -\infty & otherwise \end{cases}$$

simplify further

$$\frac{-2(\lambda_1-\lambda_2)^2 - 2(\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + \lambda_1+\lambda_2 = \frac{\lambda_1+\lambda_2+2\lambda_1\lambda_2-\lambda_1^2-\lambda_2^2}{1+\lambda_1+\lambda_2}$$

we can rewrite the Lagrange dual problem as

$$\boxed{\begin{array}{ll} \text{maximize} & \dfrac{\lambda_1+\lambda_2+2\lambda_1\lambda_2-\lambda_1^2-\lambda_2^2}{1+\lambda_1+\lambda_2} \\ \text{subject to} & \lambda_1,\lambda_2 \geq 0 \end{array}}$$

this dual is a limit, also it does not satisfy slater condition, from part (b) that KKT has no solution, hence the dual-optimal is not attained.

5.27

- primal feasibility $Gx - h \leq 0$
- dual feasibility $\nu \geq 0$
- complimentary slackness $\nu(Gx - h) = 0$
- first order condition: $2xA^TA + G\nu^T - 2A^Tb + \nu G = 0$

now we work on Lagrange:

$$L(x,\nu) = (Ax-b)^T(Ax-b) + \nu^T(GX-h) = (x^TA^T - b^T)(Ax-b) + \nu^T(GX-h)$$
$$= x^TA^TAx - bx^TA^T - b^TAx - b^Tb + \nu^TGx - \nu^Th$$
$$= x^TA^TAx + (\nu^TG - 2bA^T)x + b^Tb - \nu^Th$$

by letting $L' = (x^TA^TAx + (\nu^TG - 2bA^T)x + b^Tb - \nu^Th)' = 2xA^TA + G\nu^T - 2A^Tb = 0$,
we get

$$x^* = \frac{2A^Tb - G^T\nu}{2A^TA}$$

plug back in feasibility $Gx - h = 0$, we get

$$v^* = \frac{G2A^Tb - 2A^Ah}{GG^T}$$

Hence

$$x^* = \frac{2A^Tb - G^T\frac{G2A^Tb-2A^Ah}{GG^T}}{2A^TA} = \frac{2A^Tb - 2A^Tb - 2A^AG^{-1}h}{2A^TA}$$

Generalized inequalities

5.39

(a)

The solution is at $\textbf{rank}X = 1$, hence $x_{ij} = x_i x_j \Leftrightarrow X = xx^T$, also for for scalar, inner product/trace can be written as $\textbf{tr}(WX) = \textbf{tr}(Wxx^T) = \textbf{tr}(x^T Wx) = x^T Wx$, finally $X_{ii} = (xx^T)_{ii} = x_i^2$.
Hence the two are the same

(b)

$$\begin{array}{ll} \text{minimize} & \textbf{tr}(WX) \\ \text{subject to} & X_{ii} = 1, \forall i \\ & \textbf{rank}X = 1 \end{array}$$

here since the rank is non-convex, we need to drop the **rank** function, this 'reduction' along show that **relaxation gives a lower bound on the optimal value of TW partitioning**.
X^* **is optimal if rank X is 1**

(c)

Show the Lagrange dual of one problem is the other, since dual gives you the lower bound.

$$\boxed{\begin{array}{ll} \text{maximize} & -\mathbf{1}^T\nu \\ \text{subject to} & W + \textbf{diag}(\nu) \succeq 0 \end{array} \Longleftrightarrow \begin{array}{ll} \text{minimize} & \mathbf{1}^T\nu \\ \text{subject to} & W + \textbf{diag}(\nu) \succeq 0 \end{array}}$$

$$L = \mathbf{1}^T\nu - \textbf{tr}(X(W + \textbf{diag}(\nu))) = -\textbf{tr}(XW) + \sum_i (\nu_i - \nu_i X_{ii}) = -\textbf{tr}(WX) + \sum_i (\nu_i - \nu_i X_{ii})$$

here $\sum_i(\nu_i - \nu_i X_{ii})$ is a linear function of ν, so the coefficient $(1 - X_{ii})$ has to be zero in order to for the problem to be bounded, therefore $X_{ii} = 1$, hence the dual problem:

$$\boxed{\begin{array}{ll} \text{maximize} & -\textbf{tr}(WX) \\ \text{subject to} & X \succeq 1 \\ & X_{ii} = 1 \;\forall i \end{array} \Longleftrightarrow \begin{array}{ll} \text{minimize} & \textbf{tr}(WX) \\ \text{subject to} & X \succeq 1 \\ & X_{ii} = 1 \;\forall i \end{array}}$$

We showed (5.114) and (5.115) has same lower bound.

Additional

3.3

(a)

since $\begin{cases} \|x + 2y\| = 0 \\ \|x - y\| = 0 \end{cases} = \begin{cases} x = 0 \\ y = 0 \end{cases}$ so

```
[x + 2*y == 0, x - y == 0]
```

(b)

syntax error:

```
[square(x + y) <= u, square(u) <= x-y]
```

(c)

syntax error:

```
[inv_pos(x) + inv_pos(y) <=1]
```

(d)

syntax error:

```
[norm(hstack([u,v])) <= 3*x + y, maximum(x, 1) <= u, maximum(y,2) <=v]
```

(e)

syntax error:

```
[quad_over_lin(x+y, sqrt(y)) <= x -y +5]
```

(f)

syntax error:

```
[x >= inv_pos(y)]
```

(g)

syntax error:

```
[quad_over_lin(square(x),x) + quad_over_lin(square(y),y) <= 1]
```

(h)

```
import numpy as np
import cvxpy as cp

x = cp.Variable()
y = cp.Variable()
u = cp.Variable()
v = cp.Variable()

v_1 = np.random.randint(10, size = 10000)
v_2 = np.random.randint(10, size = 10000)
v_3 = np.random.randint(10, size = 10000)
v_4 = np.random.randint(10, size = 10000)

# constraints = [x + 2*y == 0, x - y == 0]
# constraints = [square(x + y) <= u, square(u) <= x-y]
# constraints = [inv_pos(x) + inv_pos(y) <=1]
# constraints = [norm(hstack([u,v])) <= 3*x + y, maximum(x, 1) <= u, maximum(y,2) <=v]
# constraints = [quad_over_lin(x+y, sqrt(y)) <= x -y +5]
# constraints = [x >= inv_pos(y)]
constraints  = [quad_over_lin(square(x),x) + quad_over_lin(square(y),y) <= 1]

objective = cp.Minimize(cp.sum(cp.abs(v_1 - (v_2 * x + v_3 * y + v_4 * z))))

prob = cp.Problem(objective, constraints)
print("Value of OF:", prob.solve())
print('Current value of controls:')
print(x.value, y.value)
```