.

CME307 HW2

Jacob Perricone
Stanford University
jacobp2@stanford.edu

Homework Assignment 2
Discuss Session Friday Feb 10 in Class

Optional Reading. Read Luenberger and Ye's *Linear and Nonlinear Programming 4th Edition* Chapters 4, 6, 11.

Solve the following problems:

1. Consider problem 6) of Homework Assignment 1 where the second-order cone is replaced by the p-order cone for $p \geq 1$:

$$\begin{array}{rl} \min & 2x_1 + x_2 + x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1, \\ & x_1 - \|(x_2,\ x_3)\|_p \geq 0. \end{array}$$

   (a) Write out the conic dual problem.

We have

$$\max y$$

Such That:

$$\vec{e}y + s = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

$$||(s_2, s_3)||_q \leq s_1$$

or equivalently that $s \in K^*$ a q order cone where $\frac{1}{p} + \frac{1}{q} = 1$.

   (b) Compute the dual optimal solution (y^*, s^*).

First note that

$$s = \begin{bmatrix} 2 - y \\ 1 - y \\ 1 - 6 \end{bmatrix}$$

Using the norm condition and recalling that $||x||_q = [\sum_i |x|^q]^{\frac{1}{q}}$

$$\max y$$

Such That:

$$2 - y \geq \left[2|1 - y|^q\right]^{\frac{1}{q}}$$

Since $y = 1$ is dual feasible, it follows that $y^* \geq 1$. Thus we can write

$$2 - y \geq 2^{\frac{1}{q}}[y-1]$$
$$2 \geq 2^{\frac{1}{q}}[y-1] + y$$
$$2 + 2^{\frac{1}{q}} \geq (2^{\frac{1}{q}} + 1)y$$
$$y \leq \frac{2 + 2^{\frac{1}{q}}}{1 + 2^{\frac{1}{q}}}$$

Therefore, $y^* = \frac{2+2^{\frac{1}{q}}}{1+2^{\frac{1}{q}}}$. and

$$s^* = \begin{bmatrix} 2 - y^* \\ 1 - y^* \\ 1 - y^* \end{bmatrix}$$

(c) Using the zero duality condition to compute the primal optimal solution x^*.

Since we know that the duality gap is zero we have that

$$c^T x = b^T y$$

and that

$$x^T s = 0$$

Since $s_2 = s_3 < 0$, it is apparent that $x_2 = x_3$ and thus that

$$x_1 s_1 = -2x_2 s_2$$
$$x_1(2 - y^*) = 2x_2(y^* - 1)$$
$$x_1 = 2x_2 \frac{(y^* - 1)}{2 - y^*}$$

Now using the fact taht $c^T x = b^y$ we have:

$$2(x_1 + x_2) = y^*$$
$$x_1 + x_2 = \frac{y^*}{2}$$
$$2x_2 \frac{(y^* - 1)}{2 - y^*} + x_2 = \frac{y^*}{2}$$
$$x_2[\frac{y^*}{2 - y^*}] = \frac{y^*}{2}$$
$$x_2 = \frac{2 - y^*}{2}$$

Thus $x_1 = 2\frac{(y^*-1)}{2-y^*}\frac{2-y^*}{2} = y^* - 1$.
In sumamry:

$$x = \begin{bmatrix} y^* - 1 \\ \frac{2-y^*}{2} \\ \frac{2-y^*}{2} \end{bmatrix} = \begin{bmatrix} \frac{2+2^{\frac{1}{q}}}{1+2^{\frac{1}{q}}} - 1 \\ \frac{2 - \frac{2+2^{\frac{1}{q}}}{1+2^{\frac{1}{q}}}}{2} \\ \frac{2 - \frac{2+2^{\frac{1}{q}}}{1+2^{\frac{1}{q}}}}{2} \end{bmatrix}$$

checking the constraint that $x_1 + x_2 + x_3 = 1$, we have $x_1 + 2x_2 = 2 - y^* + y^* - 1 = 1$.

2. Consider the SOCP relaxation in problem 9) of Homework Assignment 1:

$$\begin{array}{ll} \min & 0^T x \\ \text{s.t.} & \|x - a_i\|^2 \leq d_i^2,\ i = 1, 2, 3. \end{array}$$

(a) Write down the KKT optimality conditions.

Let

$$\vec{g}(x) = \begin{bmatrix} ||x-a_1||^2 - d_i^2 \\ ||x-a_2||^2 - d_2^2 \\ ||x-a_3||^2 - d_3^2 \end{bmatrix}$$

Then

$$\nabla \vec{g}(x) = \begin{bmatrix} 2(x-a_1) \\ 2(x-a_2) \\ 2(x-a_3) \end{bmatrix}$$

Let x^* be a relative minimum and suppose that it is a regular point for the constraints, then the KKT conditions require that $\exists\ \mu \in \mathbb{R}^2$ such that

$$\begin{aligned} \mu^T \nabla \vec{g}(x^*) &= \vec{0} \\ \mu^T \vec{g}(x^*) &= \vec{0} \\ \mu_T &\geq 0 \end{aligned}$$

(b) Then explain/interpret the three optimal multipliers when the true position of the sensor is inside the convex hull of the three anchors.

When the true position of the sensor is within the convex hull, we are able to find μ such that the sum of the gradient vectors are equal to zero. One can think about it as finding the forces such that they balance to zero. Within the convex hull, there is a linear combination of the multipliers such that there sum is zero.

(c) Could the true position $\bar{x} \in R^2$ of the sensor satisfy the optimality conditions if it is outside the convex hull of the three anchors? What would be the multiplier values?

When the true position is outside of the convex hull, there is no way to balance the gradient of the constraints, thus it is impossible to find a $\mu, \mu \neq \mathbf{0}$ such that $\mu^T \nabla \vec{g}(x^*) = \vec{0}$, thus all the multipliers will be zero.

3. Consider the SDP relaxation in problem 9) of Homework Assignment 1:

$$\begin{array}{lll} \max & 0 \bullet Z & \\ \text{s.t.} & (1;0;0)(1;0;0)^T \bullet Z & = 1, \\ & (0;1;0)(0;1;0)^T \bullet Z & = 1, \\ & (1;1;0)(1;1;0)^T \bullet Z & = 2, \\ & (a_i;-1)(a_i;-1)^T \bullet Z & = d_i^2,\ i = 1,2,3, \\ & Z & \succeq 0. \end{array}$$

(a) Write out the SDP dual problem, especially the dual slack matrix $U \in S^3$.

Let us define

$$w := [w_1 \quad w_2 \quad w_3]$$

and

$$\hat{w} := [\hat{w}_1 \quad \hat{w}_2 \quad \hat{w}_3]$$

. Furthermore let

$$\hat{d} = \begin{bmatrix} d_1^2 \\ d_2^2 \\ d_3^2 \end{bmatrix}$$

$$b = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

Now let $a_{i,j}$ be the jth element of the vector a_i The dual problem can be written as

$$\text{minimize } w \cdot b + \hat{w} \cdot \hat{d}$$
$$\text{such that}$$
$$w_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + w_2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + w_3 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \sum_{i=1}^{3} \hat{w}_i \begin{bmatrix} a_{i,1}^2 & a_{i,1}a_{i,2} & -a_{i,1} \\ a_{i,1}a_{i,2} & a_{i,2}^2 & -a_{i,2} \\ -a_{i,1} & -a_{i,2} & 1 \end{bmatrix} \succeq 0$$

Which is equivalent to:

$$\text{minimize } w \cdot b + \hat{w} \cdot \hat{d}$$
$$\text{such that}$$
$$\begin{bmatrix} w_1 + w_3 + \sum_{i=1}^{3} \hat{w}_i a_{i,1}^2 & w_3 + \sum_{i=1}^{3} \hat{w}_i a_{i,1} a_{i,2} & -\sum_{i=1}^{3} \hat{w}_i a_{i,1} \\ w_3 + \sum_{i=1}^{3} \hat{w}_i a_{i,1} a_{i,2} & w_2 + w_3 + \sum_{i=1}^{3} \hat{w}_i a_{i,2}^2 & -\sum_{i=1}^{3} \hat{w}_i a_{i,2} \\ -\sum_{i=1}^{3} \hat{w}_i a_{i,1} & -\sum_{i=1}^{3} \hat{w}_i a_{i,2} & \sum_{i=1}^{3} \hat{w}_i \end{bmatrix} \succeq 0$$

(b) Suppose the true position of the sensor is $\bar{x} \in R^2$. Show that if

$$\bar{U} = (-\bar{x}; 1)(-\bar{x}; 1)^T,$$

then it is an optimal slack matrix.

The primal matrix $\bar{Z}$ has the form:

$$\bar{Z} = \begin{bmatrix} I & \bar{x} \\ \bar{x}^T & x^T x \end{bmatrix} = \begin{bmatrix} 1 & 0 & \bar{x_1} \\ 0 & 1 & \bar{x_2} \\ \bar{x_1} & \bar{x_2} & x^T \bar{x} \end{bmatrix}$$

Now from the complementarity condition, we know that if $\bar{Z}\bar{U} = \mathbf{0}$, that $\bar{U}$ is optimal. The matrix $\bar{U}$ has the form

$$\bar{U} = \begin{bmatrix} \bar{x_1}^2 & \bar{x_1}\bar{x_2} & -\bar{x_1} \\ \bar{x_1}\bar{x_2} & \bar{x_2}^2 & -\bar{x_2} \\ -\bar{x_1} & -\bar{x_2} 1 & \end{bmatrix}$$

Multiplying these two matrices we find that

$$\bar{Z}\bar{U} = \begin{bmatrix} 1 & 0 & \bar{x_1} \\ 0 & 1 & \bar{x_2} \\ \bar{x_1} & \bar{x_2} & x^T \bar{x} \end{bmatrix} \begin{bmatrix} \bar{x_1}^2 & \bar{x_1}\bar{x_2} & -\bar{x_1} \\ \bar{x_1}\bar{x_2} & \bar{x_2}^2 & -\bar{x_2} \\ -\bar{x_1} & -\bar{x_2} 1 & \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus proving $\bar{U}$ is optimal.

(c) When is $\bar{U} = (-\bar{x}; 1)(-\bar{x}; 1)^T$ possible and the solution is unique?

The solution is unique when the complementarity condition holds. Noting that $\bar{U}$ is symmetric, and equating the stress matrix with the optimal, we have six equations and six unknowns. The equations are

$$w_1 + w_3 + \sum_i a_{i,1}^2 \hat{w}_i = -\bar{x}_1^2$$

$$w_2 + w_3 + \sum_i a_{i,2}^2 \hat{w}_i = -\bar{x}_1^2$$

$$w_3 + \sum_i a_{i,1} a_{i,2} \hat{w}_i = -\bar{x}_1 \bar{x}_2$$

$$\sum_i \hat{w}_i a_{i,1} = -\bar{x}_1$$

$$\sum_i \hat{w}_i a_{i,2} = -\bar{x}_2$$

$$\sum_i \hat{w}_i = -1$$

Thus when the matrix

$$W = \begin{bmatrix} 1 & 0 & 1 & a_{1,1}^2 & a_{2,1}^2 & a_{3,1}^2 \\ 0 & 1 & 1 & a_{1,2}^2 & a_{2,2}^2 & a_{3,2}^2 \\ 0 & 0 & 1 & a_{1,2}a_{1,1} & a_{2,2}a_{2,1} & a_{3,2}a_{3,1} \\ 0 & 0 & 0 & a_{1,1} & a_{2,1} & a_{3,1} \\ 0 & 0 & 0 & a_{1,2} & a_{2,2} & a_{3,2} \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} b = \begin{bmatrix} -\bar{x}_1^2 \\ -\bar{x_1}^2 \\ -\bar{x_1}\bar{x_2} \\ -\bar{x_1} \\ -\bar{x_2} \\ -1 \end{bmatrix}$$

W is of full rank, and $W \begin{bmatrix} w^T \\ \hat{w}^T \end{bmatrix} = b$ is the solution optimal and unique.

(d) Then explain/interpret the three optimal multipliers corresponding to the three distance equality constraints.

The three multipliers corresponding to the distance inequalities are the internal tensional force on edge i, j of the network. The three optimal multipliers are such that all internal forces are balanced at every sensor point.

---

4. Consider convex cone (Lecture Note 4, slide 18)

$$C = \{(t; x) : \ t > 0, \ tc(x/t) \leq 0, \ x \in R^2\},$$

where $c(x) \in R$ is a convex function. Construct the dual cone of C for each of the following $c(x)$:

(a) $c(x) = x_1^2 + 2x_2^2 - 2x_1x_2 - 1$ (ellipsoidal cone).

We have that $c(x) = x^T Q x - 1$, such that $x^T Q x = x_1^2 + 2x_2^2 - 2x_1x_2$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} q_1 & q_2 \\ q_3 & q_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 2x_2^2 - 2x_1x_2$$

$$q_1 x_1^2 + (q_2 + q_3) x_1 x_2 + q_4 x_2^2 = x_1^2 + 2x_2^2 - 2x_1x_2$$

$$q_1 = 1$$

$$q_4 = 2$$

$$q_2 + q_3 = -2$$

Thus we have

$$\phi(s) = \inf_x s^T x$$

such that:

$$x^T Q x \leq 1$$

which has optimal solution $x = \frac{-1Q^{-1}}{||Q^{-1}s||}$. Inverting this quantity, it follows that $\phi(s) = -\sqrt{s^T Q^{-1} s}$. Yielding the dual cone

$$\mathbf{K}^* = \{(\kappa, s) : \kappa - \sqrt{s^T Q^{-1} s} \geq 0\}$$

(b) $c(x) = e^{x_1} + e^{x_2} - 1$ (exponential cone).

Consider $c(x) = \sum_i e^{x_i} - 1$. It follows that

$$\phi(s) = \inf_x s^T x = \sum_i s_i x_i$$

Such That

$$\sum_i e^{x_i} < 1$$

Notice that if $s_i > 0$ then $\phi(s)$ is unbounded below, so we must have that $s_i \leq 0$ so

$$\mathbf{K}^* = \{(s_1, s_2) : s_1, s_2 \leq 0\}$$

(c) (Optional) Prove that $tc(x/t)$ is a convex function if $c(.)$ is.

Now let $g(t, x) = tc(\frac{x}{t})$. We need to prove $g(t, x)$ is convex

$$\begin{aligned}
g(\lambda_1(x_1, t_1) + \lambda_2(x_2, t_2)) &= [\lambda_1 t_1 + \lambda_2 t_2] c[\frac{\lambda_1 x_1 + \lambda_2 x_2}{\lambda_1 t_1 + \lambda_2 t_2}] \\
&= [\lambda_1 t_1 + \lambda_2 t_2] c[\frac{\lambda_1 \frac{x_1}{t_1} t_1 + \lambda_2 \frac{x_2}{t_2} t_2}{\lambda_1 t_1 + \lambda_2 t_2}] \\
&\leq [\lambda_1 t_1 + \lambda_2 t_2] \left[\frac{\lambda_1 t_1}{\lambda_1 t_1 + \lambda_2 t_2} f(\frac{x_1}{t_1}) + \frac{\lambda_2 t_2}{\lambda_1 t_1 + \lambda_2 t_2} f(\frac{x_2}{t_2}) \right] \\
&= \lambda_1 t_1 f(\frac{x_1}{t_1}) + \lambda_2 t_2 f(\frac{x_2}{t_2}) \\
&= \lambda_1 g(x_1, t_1) + \lambda_2 g(x_2, t_2)
\end{aligned}$$

Thus the proof is complete.

---

5. Consider the following parametric QCQP problem for a parameter $\kappa > 0$:

$$\begin{aligned}
\min \quad & (x_1 - 1)^2 + x_2^2 \\
\text{s.t.} \quad & -x_1 + \tfrac{x_2^2}{\kappa} \geq 0
\end{aligned} \tag{1}$$

(a) Is $x = 0$ a first-order KKT solution?

Writing out the Lagrangian for the function above we have

$$L(x_1, x_2, \lambda) = (x_1 - 1)^2 + x_2^2 + \lambda[x_1 - \frac{x_2^2}{\kappa}]$$

with $\lambda \geq 0$. Now the KKT conditions for the Lagrangian above are:

$$2(x_1 - 1) + \lambda_1 = 0$$

$$2x_2 - \frac{2\lambda_1 x_2}{\kappa} = 0$$

$$\lambda_1[x_1 - \frac{x_2^2}{\kappa}] = 0$$

If $x_1 = x_2 = 0$, for the KKT conditions to hold we must have that $\lambda_1 = 2$

(b) Is $x = 0$ a second-order KKT solution for some value of κ?

The second order conditions require the hessian to be positive definite on the tangent space. The tangent space is defined such that $T := \{y : y_1 = 0\}$, thus the subspace spanned by the second column of hessian is important.

$$\nabla_x^2(0,2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - \frac{4}{\kappa} \end{bmatrix}$$

Thus to find the eigenvalues of the matrix we have

$$det(\nabla_x^2(0,2) - \lambda I) = (2 - \lambda_1)(2 - \frac{4}{\kappa} - \lambda_2) = 0$$

Examining λ_2 it must hold that $\lambda_2 > 0$, Thus e have that $2 - \frac{4}{\kappa} > 0 \to 2 > \frac{4}{\kappa}$. Thus, $\kappa > 2$.

6. Consider the SVM problem described in Lecture Note #2:

$$\begin{array}{rl} \min & \beta + \mu\|x\|^2 \\ \text{s.t.} & a_i^T x + x_0 + \beta \geq 1,\ \forall i, \\ & b_j^T x + x_0 - \beta \leq -1,\ \forall j, \\ & \beta \geq 0. \end{array}$$

(a) Write out the Lagrangian dual of the SVM problem.

The lagrangian of the function above is

$$L(x, x_0, \beta, \lambda^a, \lambda^b, \lambda^\beta) = \beta + \mu||x||^2 - \sum_i \lambda_i^a(a_i^T x + x_0 + \beta - 1) - \sum_j \lambda_j^b(b_j^T x + x_0 - \beta + 1) - \lambda^\beta \beta$$

$$y_i^a \geq \ \forall i$$

$$y_j^b \leq \ \forall j$$

$$y^\beta \geq 0$$

The dual is defined as :

$$g(\lambda^a, \lambda^b, \lambda^\beta) = \inf_{x,\beta,x_0} \{L(x, x_0, \beta, \lambda^a, \lambda^b, \lambda^\beta)\}$$

Differentiating with respect to x, x_0, β and setting them to zero we have

$$\nabla_\beta L(x, x_0, \beta, \lambda^a, \lambda^b, \lambda^\beta) = 1 - \sum_i \lambda_i^a - \sum_j \lambda_j^b - \lambda^\beta = 0$$

$$\lambda^\beta = 1 - \sum_i \lambda_i^a - \sum_j \lambda_j^b$$

$$\nabla_x L(x, x_0, \beta, \lambda^a, \lambda^b, \lambda^\beta) = 2\mu x - \sum_i \lambda_i^a a_i - \sum_j \lambda_j^b b_j = 0$$

$$x = \frac{\sum_i \lambda_i^a a_i + \sum_j \lambda_j^b b_j}{2\mu}$$

$$\nabla_{x_0} L(x, x_0, \beta, \lambda^a, \lambda^b, \lambda^\beta) = -\sum_i \lambda_i^a - \sum_j \lambda_j^b = 0$$

$$\sum_i \lambda_i^a = -\sum_j \lambda_j^b$$

Plugging the optimal values back into g we find

$$g(\lambda^a, \lambda^b, \lambda^\beta) = \beta + \mu||x||^2 - x_0(\sum_i \lambda_i^a + \sum_j \lambda_j^b) +$$
$$\sum_i \lambda_i^a - \sum_j \lambda_j^b - \beta(\sum_i \lambda_i^a - \sum_j \lambda_j^b) - (\sum_j \lambda_j^b b_j^T + \sum_i \lambda_i^a a_i^T)x$$

Notice that the terms in the x_0 cancel and that

$$0 = \beta[\sum_j \lambda_j^b - \sum_i \lambda_i^a - \sum_j \lambda_j^b + \sum_i \lambda_i^a - 1 + 1]$$

Thus

$$\begin{aligned} g(\lambda^a, \lambda^b, \lambda^\beta) &= \sum_i \lambda_i^a - \sum_j \lambda_j^b + ux^T x - (\sum_j \lambda_j^b b_j^T + \sum_i \lambda_i^a a_i^T)x \\ &= (1 - \lambda^\beta) + \mu x^T x - 2x^T x\mu \\ &= (1 - \lambda^\beta) - \frac{1}{4\mu}||\sum_i \lambda_i^a a^i + \sum_i \lambda_j b^j||^2 \end{aligned}$$

(b) Suppose we have 6 training solution in R^2: $a_1 = (0; 0),\ a_2 = (1; 0),\ a_3 = (0; 1)$ and $b_1 = (0; 0),\ b_2 = (-1; 0),\ b_3 = (0; -1)$. Using the optimality conditions to find optimal solutions for $\mu = 0$ and $\mu = 10^{-5}$, respectively. Are the two optimal solutions unique for the given μ?

I am slightly confused by what's being asked here, but I'll give it a shot. First lets look at it mathematically, when $\mu = 0$. When $\mu = 0$ the primal optimization problem becomes:

$$\begin{array}{ll} \min & \beta \\ \text{s.t.} & a_i^T x + x_0 + \beta \geq 1,\ \forall i, \\ & b_j^T x + x_0 - \beta \leq -1,\ \forall j, \\ & \beta \geq 0. \end{array}$$

For our case lets break down the constraints set, we have

$$\begin{aligned} &\min \beta \\ &\text{such that:} \\ x_0 + \beta &\geq 1 \\ x_1 + x_0 + \beta &\geq 1 \\ x_2 + x_0 + \beta &\geq 1 \\ x_0 - \beta &\leq -1 \\ -x_1 + x_0 - \beta &\leq -1 \\ -x_2 + x_0 - \beta &\leq -1 \end{aligned}$$

Now from the inequalities we see that $x_0 + \beta \geq 1$ and $x_0 - \beta \leq -1$. We want β as small as possible and this to hold. It is fairly clear that beta attains a minimum while satisfying the constraint when β is close to one and x_0 is close to zero, so we know that $x_0 \approx 0, \beta \approx 1$.

Thus it must follow that $x_1, x_2 \geq 0$. Without the norm term, the SVM aims to find a separating hyperplane but does not care about minimizing the norm of the values, i.e. maximizing the margins, thus with these constraints we can see that there x_1, x_2 values are not unique since they only require $x_1, x_2 \geq 0$. Intuitevly since we have eliminated the term with the norm, the SVM is just finding a separating hyperplane and not one that maximizes the margins. So it could draw one with $x_0 = 0, \beta = 1, x_1 = 0, x_2 = 0$, or $x_0 = 0, \beta = 1, x_1 = 0, x_2 = 10$ or $x_0 = 0, \beta = 1, x_1 = 3, x_2 = 3$.

When $\mu > 0$, however, the optimization should find values that minimize the norm, i.e. ones where $x_1 = x_2$. My instinct says that it should set $x_1, x_2 = 0$, but from the numerical experiments it only does this when $\mu >> 0$. For varying μ one can observe a trend in the hyperplane parameterization. As μ

increases the hyperplane goes from $\mu_0 = (4.7821, 4.7821) \rightarrow \mu_{10^{-5}} = (0.01851330, 0.01851330) \rightarrow \mu_{1000} = (3.76e^{-06}, 3.76e^{-06})$, indicating that it penalizes for the norm of the parameters. As the penalty μ gets higher, the unique minimizing norm$x_1 = x_2 = 0$, is chosen.

```
clear all
close all

% Problem 6(b)
mu = [0, 10^-5, 1, 2, 10, 100,100];
a = [0, 1, 0; 0 0, 1];
b = [0,-1, 0; 0, 0,-1];


%% Find optimal hyperplane for different Beta
for j = 1:size(mu,2)
    cvx_begin quiet
     variable B
     variable x1
     variable x2
     variable x0
    mu(j)
    opt = B + mu(j)*power(x1,2) + mu(j)*power(x2,2);

    minimize(opt);
     subject to

    B >= 0;


    x0 + B >=1
    x1 + x0 + B >= 1
    x2 + x0 + B >=1

    x0 - B <= -1
    -x1 + x0 - B <= -1
    -x2 + x0 - B <= -1

    B >= 0
    cvx_end

    ex1{j} = x1;
    ex2{j} = x2;
    ex{j} = [x1, x2];e
    ex0{j} = x0;

    eB{j} = B;
end
```

---

7. Find the Lagrangian dual of the barrier optimization problem where given parameter $\mu > 0$:

$$\begin{array}{lll} \min & c^T x - \mu \sum_{j=1}^{n} \ln(x_j), & \\ \text{s.t.} & Ax & = b \\ & x & > 0, \end{array}$$

where we assume that the problem has an interior-point solution, and there exists a vector $y \in R^m$ such that $c - A^T y > 0$.

What are the first-order KKT optimality conditions?

The Lagrangian of the problem above is

$$L(x,y) = \sum_j (c_j - a_j^T y)x_j - \mu \sum_{j=1}^{n} \ln(x_j) + b^T y$$

Now let us define the operator $x^{\div}$ as the element wise reciprocal operator such that $x_i^{\div} = \frac{1}{x_i}$ for all i. Notice that we don't need a multiplier for x since a negative x would drive the objective to infinity. Also notice that it must hold that $(c_j - a_j^t y) > 0 \; \forall j$ or the Lagrangian is unbounded from below.
Now we have

$$g(y) = \inf_x L(x,y)$$

Calculating the gradient with respect to x

$$\nabla_x L(x,y) = c - \mu x^{\div} - y^T A = 0$$
$$\frac{1}{\mu}[c - y^T A] = x^{\div}$$
$$x_i = \frac{\mu}{c_i - a_j^T y}$$

Thus:

$$g(y) = \sum_j (c_j - a_j^T y)\frac{\mu}{c_j - a_j^T} - \mu \sum_{j=1}^{n} \ln[\frac{\mu}{c_j - a_j^T y}] + b^T y \quad = \sum_j \mu[1 - \ln[\frac{u}{c_j - a_j^T y}]] + b^T y$$

The KKT conditions of the primal problem are:

$$(1)(c_j - a_j^T y) - \frac{\mu}{x_j} = 0 \text{ for all } j$$
$$(2) y^T(Ax - b) = 0$$

---

8. Consider a generalized Arrow–Debreu equilibrium problem in which the market has n agents and m goods. Agent i, $i = 1, ..., n$, has a bundle amount of $w_i = (w_{i1}, w_{i2}, \ldots, w_{im}) \in R^m$ goods initially and has a linear utility function whose coefficients are $u_i = (u_{i1}, u_{i2}, \ldots, u_{im}) > 0 \in R^m$. The goal is to price each good so that the market clears. Note that, given the price vector $p = (p_1, p_2, \ldots, p_m) > 0$, agent i's utility maximization problem is:

$$\begin{array}{ll} \text{maximize} & u_i^T x_i \\ \text{subject to} & p^T x_i \le p^T w_i \\ & x_i \ge 0 \end{array}$$

(a) For a given $p \in R^m$, write down the optimality conditions for agent i's utility maximization problem. Without loss of generality, you may fixed $p_m = 1$ since the budget constraint are homogeneous in p.

The lagrangian for agent i's utility function is

$$L(x, \lambda_1, \lambda_2) = u_i^T x_i + \lambda_1[p^T x_i - p^T w_i] - \lambda_2^T x_i$$

with $\lambda_1, \lambda_2 \ge 0$. The optimality conditions are thus

$$(1) u_i + \lambda_1 p - \lambda_2 = 0$$
$$(2) \lambda_1(p^T x_i - p^T w_i) = 0$$
$$(3) \lambda_2^T x_i = 0$$

(b) (5pts.) Suppose that $p \in R^m$ and $x_i \in R^m$ satisfy the constraints:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} w_i$$

$$\frac{u_i^T x_i}{p^T w_i} p_j \geq u_{ij} \qquad \forall i, j$$

$$x_{ij} \geq 0, p_j \geq 0 \qquad \forall i, j$$

for $i = 1, \ldots, n$. Show that p is then an equilibrium price vector.

We must show that $p^T w_i = p^t x_i \; \forall i$ Multiplying both sides of constraint two by the sum of x_{ij} over j we have

$$\begin{aligned}
\frac{u_i^T x_i}{p^T w_i} \sum_j p_j x_{ij} &\geq \sum_j u_{ij} x_{ij} \\
\frac{u_i^T x_i}{p^T w_i} \sum_j p_j x_{ij} &\geq u_i^T x_i \\
\frac{1}{p^T w_i} \sum_j p_j x_{ij} &\geq 1 \\
\sum_j p_j x_{ij} &\geq p^T w_i \\
p^T x_i &\geq p^T w_i
\end{aligned}$$

Now from the first constraint we have

$$\sum_i w_i = \sum_i x_i$$

Since $p \geq 0$

$$\begin{aligned}
\sum_i w_i &= \sum_i x_i \\
p^T \sum_i w_i &= p^T \sum_i x_i \\
\sum_i p^T w_i &= \sum_i p^T x_i
\end{aligned}$$

Now since we know that each element of $p^T x_i \geq p^T w_i$, the only way for the equality to hold is if $p^T w_i = p^T x_i \; \forall i$.

(c) For simplicity, assume that all u_{ij} are positive so that all p_j are positive. By introducing new variables $y_j = \log(p_j)$ for $j = 1, ..., m$, the conditions can be written as follows:

$$\begin{array}{ll}
\min & 0 \\
\text{s.t.} & \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} w_i \\
& \log(u_i^T x_i) - \log\left(\sum_{k=1}^{m} w_{ik} e^{y_k}\right) + y_j \geq \log(u_{ij}) \quad \forall i, j \\
& x_{ij} \geq 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall i, j
\end{array}$$

Show that this problem is convex in x_{ij} and y_j. (Hint: Use the fact that $\log\left(\sum_{k=1}^{m} w_{ik} e^{y_k}\right)$ is a convex function in the y_k's.)

First notice that the first constraint set $C_1 := \{x_i : \sum_{i=1}^m x_i - \sum_{i=1}^n w_i = 0\}$ is a hyperplane in $x_{i,j}$, and is thus convex. Now the second constraint $C_2 := \{(x_i, y_j) \log(u_{ij}) - y_j - \log(\sum_j u_{ij} x_{ij}) + \log[\sum_k w_{ik} e^{y_k}]\}$. The constraint set is a linear combination of two convex functions of $y_j \to (-y_j, \log[\sum_k w_{ik} e^{y_k}])$, and thus is convex in y_j. Now for x_{ij}, note that the function $\log \frac{1}{x}$ is convex in x for $0 \leq x \leq \infty$. So, $-\log(\sum_j u_{ij} x_{ij}) = \log[\frac{1}{\sum_j u_{ij} x_{ij}}]$, which is convex in x_{ij}. Since the constraints set is the intersection of two convex sets of x_{ij}, y_j, the problem is convex in x_{ij}, y_j.

(d) Consider the Fisher example on slide 23 of Lecture Note #5 with two agents and two goods, where there is no fixed budgets. Rather, let

$$w_1 = (1;\ 0) \quad \text{and} \quad w_2 = (0;\ 1)$$

that is, agent 1 brings in one unit good x and agent brings in one unit of good y. Find the Arrow–Debreu equilibrium prices, where you may assume $p_y = 1$.
Solving this in CVX yields that the price is equal to 2.

---

9. (20pts) Computation Team Work: See the attached report!