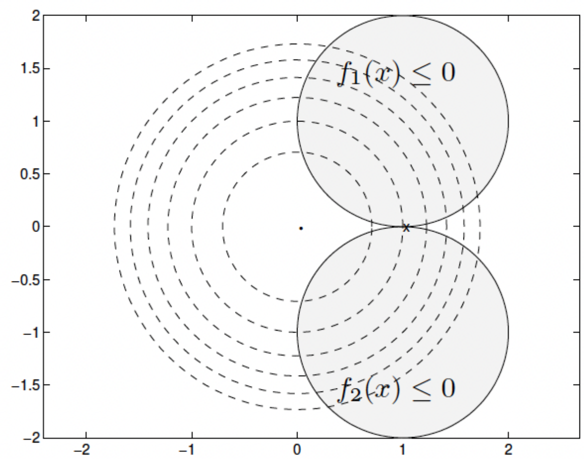
DSA3102: Solutions to Tutorial Set 10
Assigned: 19/10/23

1. BV Problem 5.26

   Solutions:

   (a) The figure shows the feasible set (the intersection of the two shaded disks) and some contour lines of the objective function. There is only one feasible point, $(1, 0)$, so it is optimal for the primal problem, and we have $p^* = 1$.



   (b) The KKT conditions are

$$
\begin{aligned}
(x_1 - 1)^2 + (x_2 - 1)^2 &\leq 1 \\
(x_1 - 1)^2 + (x_2 + 1)^2 &\leq 1 \\
\lambda_1 &\geq 0 \\
\lambda_2 &\geq 0 \\
2x_1 + 2\lambda_1(x_1 - 1) + 2\lambda_2(x_1 - 1) &= 0 \\
2x_2 + 2\lambda_1(x_2 - 1) + 2\lambda_2(x_2 + 1) &= 0 \\
\lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) = \lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1) &= 0
\end{aligned}
$$

   At $x = (1, 0)$, these conditions reduce to

$$\lambda_1 \geq 0, \qquad \lambda_2 \geq 0, \qquad 2 = 0, \qquad -2\lambda_1 + 2\lambda_2 = 0,$$

   which (clearly, in view of the third equation) have no solution.

(c) The Lagrange dual function is given by

$$g(\lambda_1, \lambda_2) = \inf_{x_1, x_2} L(x_1, x_2, \lambda_1, \lambda_2)$$

where

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

L reaches its minimum at

$$x_1 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2} \qquad x_2 = \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2},$$

and we find that

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

where we interpret $a/0 = 0$ if $a = 0$ and as $-\infty$ if $a < 0$. The Lagrange dual problem is given by

$$\max \frac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \qquad \lambda_1, \lambda_2 \geq 0$$

Since g is symmetric, the optimum (if it exists) occurs with $\lambda_1 = \lambda_2$. The dual function then simplifies to

$$g(\lambda_1, \lambda_2) = \frac{2\lambda_1}{2\lambda_1 + 1}$$

We see that $g(\lambda_1, \lambda_2)$ tends to 1 as $\lambda_1 \to \infty$. We have $d^* = p^* = 1$, but the dual optimum is not attained.

Recall that the KKT conditions only hold if (1) strong duality holds, (2) the primal optimum is attained, and (3) the dual optimum is attained. In this example, the KKT conditions fail because the dual optimum is not attained.

2. BV Problem 5.27

Solution:

(a) The Lagrangian is

$$L(x, \nu) = \|Ax - b\|_2^2 + \nu^T(Gx - h)$$

with minimizer

$$x = -\frac{1}{2}(A^TA)^{-1}(G^T\nu - 2A^Tb)$$

Plugging this into the Lagrangian gives the dual function

$$g(\nu) = -\frac{1}{4}(G^T\nu - 2A^Tb)^T(A^TA)^{-1}(G^T\nu - 2A^Tb) - \nu^Th + b^Tb.$$

(b) The optimality conditions are

$$2A^T(Ax^* - b) + G^T\nu^* = 0, \quad Gx^* = h.$$

(c) From the first equation,

$$x^* = (A^TA)^{-1}(A^Tb - (1/2)G^T\nu^*).$$

Plugging this into the second equation yields

$$G(A^TA)^{-1}A^Tb - (1/2)G(A^TA)^{-1}G^T\nu^* = h$$

i.e.,

$$\nu^* = -2(G(A^TA)^{-1}G^T)^{-1}(h - G(A^TA)^{-1}A^Tb)$$

Substituting in the first expression gives an analytical expression for x^*.

3. Consider the problem

$$\min\ f(x), \qquad \text{s.t.} \qquad 5x_1 + x_2 \geq 4,$$

where

$$f(x) = \begin{cases} 10x_1 + 3x_2 & (x_1, x_2) \in \{0,1\}^2 \\ +\infty & \text{else} \end{cases}$$

(a) Sketch the set of constraint-cost pairs

$$\{(4 - 5x_1 - x_2, 10x_1 + 3x_2) : x_1, x_2 \in \{0,1\}\}$$

Solution: This is obvious.

(b) Sketch the dual function.

Solution: The Lagrangian is

$$L(x,\mu) = 10x_1 + 3x_2 + \mu(4 - 5x_1 - x_2)$$

and the dual function is

$$g(\mu) = \inf_{x_1,x_2\in\{0,1\}} \{4\mu + (10 - 5\mu)x_1 + (3-\mu)x_2\} = \begin{cases} 4\mu & \mu \in [0,2] \\ 10 - \mu & \mu \in [2,3] \\ 13 - 2\mu & \mu \in [3,\infty) \end{cases}$$

(c) Solve the problem and its dual.

Solution: By inspection, we see that $x^* = (1,0)$ and $p^* = 10$. From the dual, we see that $d^* = 8$. Thus, there is a duality gap of $p^* - d^* = 2$.

4. **Solution:** A straightforward calculation yields the dual function

$$g(\lambda) = \min_{x\in\mathbb{R}^n} \left\{\|z - x\|_2^2 + \lambda^TAx\right\} = -\frac{\|A^T\lambda\|_2^2}{4} + \lambda^TAz$$

Thus the dual problem is equivalent to

$$\min_{\lambda\in\mathbb{R}^m} \left\{\frac{\|A^T\lambda\|_2^2}{4} - \lambda^TAz + \|z\|_2^2\right\}$$

or

$$\min_{\lambda\in\mathbb{R}^m} \left\|z - \frac{A^T\lambda}{2}\right\|_2^2$$

This is the problem of projecting z on the subspace spanned by the rows of A.

5. **Solutions**:

(a) Lagrange optimality yields

$$\nabla f(x^*) + \nu^* \nabla h(x^*) = 0$$

which is

$$2x^* + (\nu^*, \dots, \nu^*)^T = 0$$

Hence

$$x^* = -\frac{1}{2}(\nu^*, \dots, \nu^*)^T$$

But the sum of the vector x^* must be one by primal optimality so

$$x^* = (1/n, \dots, 1/n)^T.$$

Problem is convex so x^* is globally optimal.

(b) Lagrange optimality yields

$$(1, 1, \dots, 1)^T + 2\nu^* x^* = 0$$

So

$$x^* = \left(-\frac{1}{2\nu^*}, \dots, -\frac{1}{2\nu^*}\right)^T$$

Furthermore the norm of x^* must equal 1 so

$$x^* = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)^T \quad \text{or} \quad x^* = -\left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)^T$$

Furthermore,

$$\nabla^2 f(x^*) + \nu^* \nabla^2 h(x^*) = 2\lambda^* I = -\sqrt{n} I \quad \text{or} \quad \sqrt{n} I$$

So the only local minimum is $x^* = -(1, 1, \dots, 1)^T/\sqrt{n}$ so it is the global minimum.

(c) Since Q is positive definite, it admits the eigendecomposition

$$Q = \sum_{i=1}^{n} \lambda_i q_i q_i^T$$

where we can assume, without loss of generality, that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$. The Lagrangian is

$$L(x, \nu) = x^T x - \nu(x^T Q x - 1).$$

Stationarity reads

$$2x - \nu(2Qx) = 0 \qquad \Longleftrightarrow \qquad (I - \nu Q)x = 0 \qquad \Longleftrightarrow \qquad \nu Q x = x.$$

In other words, x is an eigenvector of

$$\nu Q = \sum_{i=1}^{n} (\nu \lambda_i) q_i q_i^T$$

with eigenvalue 1. Primal feasibility is

$$x^T Q x = 1.$$

To ensure that x is an eigenvector of νQ with eigenvalue 1, we set $\nu > 0$ such that $\nu \lambda_i = 1$ for some $i = 1, 2, \dots, n$. At this point, there are n possible choices for ν. At the same time, primal feasibility says that

$$x^T \frac{1}{\nu} x = 1 \qquad \Longleftrightarrow \qquad \nu = x^T x.$$

Since we want to minimize $x^T x$, we choose ν as small as possible, i.e., $\nu = 1/\lambda_1$. Then, x is simply the eigenvector corresponding to the largest eigenvalue of Q, which is positive.