Homework 6

Submit answers for problems 1–6. Problems 7–10 are practice problems.

1. Exercise T5.19.
2. Exercise A5.20.
3. Exercise A5.30.
4. Exercise A5.14.
5. Exercise T5.29.
6. Exercise T5.21(a,b,c).
7. Exercise T5.17.

Solution. The problem can be expressed as

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m, \end{array}$$

if we define $f_i(x)$ as the optimal value of the LP

$$\begin{array}{ll} \text{maximize} & x^T a \\ \text{subject to} & C_i a \preceq d_i, \end{array}$$

where a is the variable and x is treated as a problem parameter. The Lagrange dual of this LP is given by

$$\begin{array}{ll} \text{minimize} & d_i^T z \\ \text{subject to} & C_i^T z = x \\ & z \succeq 0. \end{array}$$

The optimal value of the dual LP is also equal to $f_i(x)$. Therefore $f_i(x) \leq b_i$ if and only if there exists a z_i with

$$d_i^T z_i \leq b_i, \qquad C_i^T z_i = x, \qquad z_i \succeq 0.$$

If we substitute these three constraints for the constraint $f_i(x) \leq b_i$ we obtain the LP in the problem statement.

8. Exercise A5.3.

Solution. The Lagrangian is

$$L(x, z, \mu) = \sum_k x_k \log(x_k/y_k) + b^T z - z^T A x + \mu - \mu \mathbf{1}^T x.$$

Minimizing over x_k gives the conditions

$$1 + \log(x_k/y_k) - a_k^T z - \mu = 0, \quad k = 1, \ldots, n.$$

This has a solution

$$x_k = y_k e^{a_k^T z + \mu - 1}, \quad k = 1, \ldots, n,$$

for any value of z, μ. Substituting x in L gives the Lagrange dual function

$$g(z, \mu) = b^T z + \mu - \sum_{k=1}^n y_k e^{a_k^T z + \mu - 1}$$

and the dual problem is

$$\text{maximize} \quad b^T z + \mu - \sum_{k=1}^n y_k e^{a_k^T z + \mu - 1}.$$

This can be simplified a bit if we optimize over μ by setting the derivative equal to zero:

$$\mu = 1 - \log \sum_{k=1}^n y_k e^{a_k^T z}.$$

After this simplification the dual problem reduces to the problem in the assignment.

9. Exercise A5.32. Also explain how the answers change if the Euclidean norm in the first constraint is replaced by the ℓ_1-norm, *i.e.*, for the problem

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & |x_1| + |x_2| \leq x_2 \\ & -x_1 \leq 1. \end{array}$$

Solution.

(a) The feasible set is $\{(x_1, x_2) \mid x_1 = 0, x_2 \geq 0\}$. Since $x_1 = 0$ at all feasible points, all feasible points are optimal. $p^\star = 0$.

(b) The Lagrangian is

$$\begin{array}{rcl} L(x_1, x_2, \lambda_1, \lambda_2) & = & x_1 + \lambda_1 \left(\|(x_1, x_2)\|_2 - x_2 \right) - \lambda_2 (x_1 + 1) \\ & = & (1 - \lambda_2) x_1 - \lambda_1 x_2 + \lambda_1 \|(x_1, x_2)\|_2 - \lambda_2 \\ & = & \left[\begin{array}{c} 1 - \lambda_2 \\ -\lambda_1 \end{array} \right]^T \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] + \lambda_1 \left\| \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] \right\|_2 - \lambda_2. \end{array}$$

Now we use the fact that for given a and $\gamma \geq 0$,

$$\inf_x (a^T x + \gamma \|x\|_2) = \begin{cases} 0 & \|a\|_2 \leq \gamma \\ -\infty & \text{otherwise.} \end{cases}$$

(See lecture 5, page 7.) Therefore the dual function is

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \inf_{x_1, x_2} L(x_1, x_2, \lambda_1, \lambda_2) \\ &= \begin{cases} -\lambda_2 & \|(1-\lambda_2, -\lambda_1)\|_2 \leq \lambda_1 \\ -\infty & \text{otherwise,} \end{cases} \\ &= \begin{cases} -1 & \lambda_1 \geq 0, \lambda_2 = 1 \\ -\infty & \text{otherwise,} \end{cases} \end{aligned}$$

and the dual problem is

$$\begin{array}{ll} \text{maximize} & -1 \\ \text{subject to} & \lambda_1 \geq 0. \end{array}$$

(c) The dual cost function is equal to -1 at all feasible points, so all feasible points are optimal and $d^\star = -1$.
Weak duality ($0 = p^\star \geq d^\star = -1$) holds, as it must for any optimization problem. Strong duality does not hold, but the problem does not satisfy Slater's condition.

The answers for the ℓ_1-norm constraint are as follows.

(a) The feasible set is still $\{(x_1, x_2) \mid x_1 = 0, x_2 \geq 0\}$ and $p^\star = 0$.

(b) The Lagrangian is

$$\begin{aligned} L(x_1, x_2, \lambda_1, \lambda_2) &= x_1 + \lambda_1(\|(x_1, x_2)\|_1 - x_2) - \lambda_2(x_1 + 1) \\ &= \begin{bmatrix} 1-\lambda_2 \\ -\lambda_1 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \lambda_1 \left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_1 - \lambda_2. \end{aligned}$$

Using the same result from lecture 5, and the fact that dual norm of the ℓ_1-norm is the ℓ_∞-norm, we obtain the dual function

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \begin{cases} -\lambda_2 & \|(1-\lambda_2, -\lambda_1)\|_\infty \leq \lambda_1 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} -\lambda_2 & |1-\lambda_2| \leq \lambda_1 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -\lambda_2 \\ \text{subject to} & |1-\lambda_2| \leq \lambda_1 \\ & \lambda_2 \geq 0. \end{array}$$

This can be simplified to

$$\begin{array}{ll} \text{maximize} & -\lambda_2 \\ \text{subject to} & \lambda_2 \geq 0. \end{array}$$

(c) The dual optimal value is $d^\star = 0$. Now strong duality holds, even though the problem still does not satisfy Slater's condition.

10. Excercise A5.26.

Solution.

(a) The Lagrangian is $L(x, y, z) = \|y\|_2 + \gamma\|x\|_1 + z^T(Ax - b - y)$. The infimum over x and y is

$$\inf_{x,y} L(x, y, z) = \begin{cases} -b^T z & \|z\|_2 \leq 1, \ \|A^T z\|_\infty \leq \gamma \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -b^T z \\ \text{subject to} & \|A^T z\|_\infty \leq \gamma \\ & \|z\|_2 \leq 1. \end{array}$$

(b) The KKT conditions for optimality of x, y, z are:

i. *Primal feasibility:* $y = Ax - b$.

ii. *Dual feasibility:* $\|z\|_2 \leq 1$ and $\|A^T z\|_\infty \leq \gamma$.

iii. *Minimum of Lagrangian:* x, y minimize $L(\tilde{x}, \tilde{y}, z)$. Therefore

$$\|y\|_2 - z^T y = \inf_{\tilde{y}} \left(\|\tilde{y}\|_2 - z^T \tilde{y} \right) = 0,$$

and

$$\gamma\|x\|_1 + z^T Ax = \inf_{\tilde{x}} \left(\gamma\|\tilde{x}\|_1 + z^T A\tilde{x} \right) = 0.$$

We apply these conditions to $x = x^\star$ and $y = Ax^\star - b$. The first part of condition (iii), combined with $\|z\|_2 \leq 1$ and $y \neq 0$, implies that $z = y/\|y\|_2 = r$. From condition (ii), we have $\|A^T r\|_\infty \leq \gamma$ and from the second part of condition (iii), $\gamma\|x\|_1 + r^T Ax = 0$.

(c) The condition $-(A^T r)^T x^\star = \gamma\|x^\star\|_1$ with $\|A^T r\|_\infty \leq \gamma$ holds only if, for each column a_k,

$$a_k^T r = \gamma \text{ if } x_k^\star < 0, \qquad a_k^T r = -\gamma \text{ if } x_k^\star > 0, \qquad |a_k^T r| \leq \gamma \text{ if } x_k^\star = 0.$$

Now, from the Cauchy–Schwarz inequality, since $\|r\|_2 = 1$, we have

$$|a_i^T r| \leq \|a_i\|_2 \|r\|_2 = \|a_i\|_2.$$

Therefore if $\|a_i\|_2 < \gamma$, we must have $x_i^\star = 0$.