Homework 6 solutions

1. Exercise T5.19.

Solution.

(a) For simplicity we assume that the elements of x are sorted:

$$x_1 \geq x_2 \geq \cdots \geq x_n.$$

It is easy to see that the optimal value is

$$x_1 + x_2 + \cdots + x_r$$

and obtained by choosing $y_1 = y_2 = \cdots = y_r = 1$ and $y_{r+1} = \cdots = y_n = 0$.

(b) We first change the objective from maximization to minimization:

$$\begin{array}{ll} \text{minimize} & -x^T y \\ \text{subject to} & 0 \preceq y \preceq \mathbf{1} \\ & \mathbf{1}^T y = r. \end{array}$$

We introduce a Lagrange multiplier λ for the lower bound, u for the upper bound, and t for the equality constraint. The Lagrangian is

$$\begin{aligned} L(y, \lambda, u, t) &= -x^T y - \lambda^T y + u^T (y - \mathbf{1}) + t(\mathbf{1}^T y - r) \\ &= -\mathbf{1}^T u - rt + (-x - \lambda + u + t\mathbf{1})^T y. \end{aligned}$$

Minimizing over y yields the dual function

$$g(\lambda, u, t) = \begin{cases} -\mathbf{1}^T u - rt & -x - \lambda + u + t\mathbf{1} = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is to maximize $g(\lambda, u, t)$ subject to $\lambda \succeq 0$ and $u \succeq 0$:

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T u - rt \\ \text{subject to} & -\lambda + u + t\mathbf{1} = x \\ & \lambda \succeq 0, \quad u \succeq 0, \end{array}$$

After changing the objective to minimization (*i.e.*, undoing the sign change we started with), we obtain

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T u + rt \\ \text{subject to} & u + t\mathbf{1} \succeq x \\ & u \succeq 0. \end{array}$$

We eliminated λ by noting that it acts as a slack variable in the first constraint.

(c) The problem is equivalent to the QP

$$\begin{array}{ll} \text{minimize} & x^T \Sigma x \\ \text{subject to} & \overline{p}^T x \geq r_{\min} \\ & \mathbf{1}^T x = 1, \quad x \succeq 0 \\ & \lfloor n/10 \rfloor t + \mathbf{1}^T u \leq 0.8 \\ & t\mathbf{1} + u \succeq x \\ & u \succeq 0, \end{array}$$

with variables x, u, t, v.

2. Exercise A5.20.

Solution. The Lagrangian is

$$L(x,z) = \sum_{i=1}^{n} (\phi(x_i) - x_i(a_i^T z)) + b^T z$$

where a_i is the ith column of A. The dual function is

$$\begin{aligned} g(z) &= b^T z + \sum_{i=1}^{n} \inf_{x_i} (\phi(x_i) - x_i(a_i^T z)) \\ &= b^T z + \sum_{i=1}^{n} h(a_i^T z) \end{aligned}$$

where $h(y) = \inf_u (\phi(u) - yu) = -\phi^*(y)$. With this notation, the dual problem can be written as

$$\text{maximize} \quad b^T z + \sum_{i=1}^{n} h(a_i^T z).$$

We now work out an expression for the function h. If $|y| \leq 1/c$, the minimizer in the definition of h is $u = 0$, and $h(y) = 0$. Otherwise, we find the minimum by setting the derivative equal to zero. If $y > 1/c$, we solve

$$\phi'(u) = \frac{c}{(c-u)^2} = y.$$

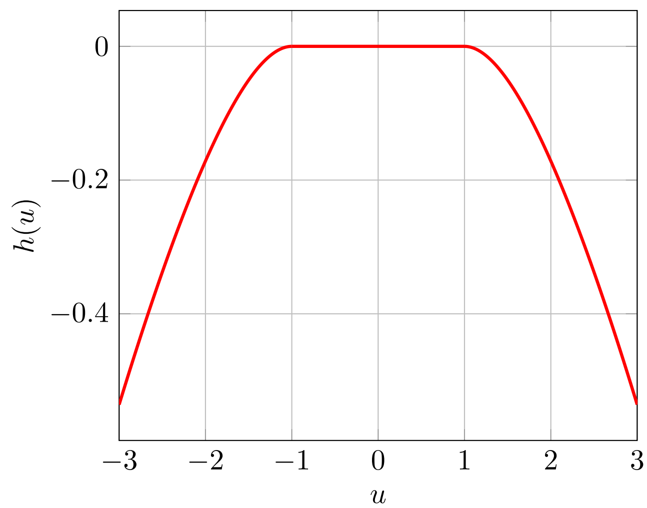
The solution is $u = c - (c/y)^{1/2}$ and we obtain $h(y) = -(1-\sqrt{cy})^2$. If $y < -1/c$, we solve

$$\phi'(u) = -\frac{c}{(c+u)^2} = y.$$

The solution is $u = -c + (-c/y)^{1/2}$ and $h(y) = -(1-\sqrt{-cy})^2$. Combining the different cases, we can write

$$h(u) = \begin{cases} -\left(1 - \sqrt{c|y|}\right)^2 & |y| > 1/c \\ 0 & \text{otherwise.} \end{cases}$$

The figure shows the function h for $c = 1$.



3. Exercise A5.30.

Solution.

(a) The Lagrangian is

$$L(x, y, z) = c^T x + \frac{1}{\mu}\sum_{i=1}^m \log(1 + e^{\mu y_i}) + z^T(Ax - b - y).$$

The minimum over x is unbounded below unless $A^T z + c = 0$. To find the minimum over y we note the function is separable. Setting the derivative with respect to y_i to zero gives

$$\frac{e^{\mu y_i}}{1 + e^{\mu y_i}} = z_i, \qquad y_i = \frac{1}{\mu}\log\frac{z_i}{1 - z_i}$$

and

$$\begin{aligned} &\inf_{y_i} \frac{1}{\mu}\log(1 + e^{\mu y_i}) - z_i y_i \\ &= \begin{cases} -(1/\mu)(z_i \log z_i + (1 - z_i)\log(1 - z_i)) & 0 \le z_i \le 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

(with the interpretation $0 \log 0 = 0$). We therefore obtain the dual

$$\begin{array}{ll} \text{maximize} & -b^T z - \frac{1}{\mu}\sum_{i=1}^m (z_i \log z_i + (1 - z_i)\log(1 - z_i)) \\ \text{subject to} & A^T z + c = 0 \\ & 0 \preceq z \preceq \mathbf{1}. \end{array}$$

(b) Plugging in the optimal primal solution $x^\star$ of the LP in the cost function (28) in the assignment gives

$$\begin{aligned} q^\star &\le c^T x^\star + \frac{1}{\mu}\sum_{i=1}^m \log(1 + e^{\mu(a_i^T x^\star - b_i)}) \\ &\le p^\star + \frac{m \log 2}{\mu} \end{aligned}$$

because $a_i^T x^\star - b_i \leq 0$. Plugging in the optimal dual solution $z^\star$ of the LP in the dual of (28) gives

$$\begin{aligned} q^\star &\geq -b^T z^\star - \frac{1}{\mu}\sum_{i=1}^m (z_i^\star \log z_i^\star + (1-z_i^\star)\log(1-z_i^\star)) \\ &\geq p^\star \end{aligned}$$

because $u \log u \leq 0$ on $[0,1]$.

4. Exercise A5.14.

Solution.

(a) The KKT conditions are

$$\frac{1}{a^T x} a + \frac{1}{b^T x} b \preceq \nu \mathbf{1} \qquad x \succeq 0, \qquad \mathbf{1}^T x = 1,$$

plus the complementary slackness conditions

$$x_k \left(\nu - \frac{1}{a^T x} a_k - \frac{1}{b^T x} b_k \right) = 0, \quad k = 1, \ldots, n.$$

We show that $x = (1/2, 0, \ldots, 0, 1/2)$, $\nu = 2$ solve these equations, and hence are primal and dual optimal.

The feasibility conditions $x \succeq 0$, $\mathbf{1}^T x = 1$ obviously hold, and the complementary slackness conditions are satisfied for $k = 2, \ldots, n-2$. It remains to verify the inequalities

$$\frac{a_k}{a^T x} + \frac{b_k}{b^T x} \leq \nu, \quad k = 1, \ldots, n, \tag{1}$$

and the complementary slackness condition

$$x_k \left(\nu - \frac{1}{a^T x} a_k - \frac{1}{b^T x} b_k \right) = 0, \quad k = 1, n. \tag{2}$$

For $x = (1/2, 0, \ldots, 0, 1/2)$, $\nu = 2$ the inequality (1) holds with equality for $k = 1$ and $k = n$, since

$$\frac{a_1}{a^T x} + \frac{b_1}{b^T x} = \frac{2a_1}{a_1 + a_n} + \frac{2/a_1}{1/a_1 + 1/a_n} = 2,$$

and

$$\frac{a_n}{a^T x} + \frac{b_n}{b^T x} = \frac{2a_n}{a_1 + a_n} + \frac{2/a_n}{1/a_1 + 1/a_n} = 2.$$

Therefore also (2) is satisfied. The remaining inequalities in (1) reduce to

$$\frac{a_k}{a^T x} + \frac{b_k}{b^T x} = 2\frac{a_k + a_1 a_n / a_k}{a_1 + a_n} \leq 2, \quad k = 2, \ldots, n-1.$$

This is valid, since it holds with equality for $k = 1$ and $k = n$, and the function $t + a_1 a_n/t$ is convex in t, so

$$\frac{t + a_1 a_n/t}{a_1 + a_n} \leq 2$$

for all $t \in [a_n, a_1]$.

(b) Diagonalize A using its eigenvalue decomposition $A = Q\Lambda Q^T$, and define $a_k = \lambda_k$, $b_k = 1/\lambda_k$, $x_k = (Q^T u)_k^2$. From part (a), $Q^T u = (1/\sqrt{2}, 0, \ldots, 1/\sqrt{2})$ is optimal. Therefore,

$$\begin{aligned} (u^T A u)(u^T A^{-1} u) &\leq \frac{1}{4}(\lambda_1 + \lambda_n)(\lambda_1^{-1} + \lambda_n^{-1}) \\ &= \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2. \end{aligned}$$

5. Exercise T5.29.

Solution. The Lagrangian is

$$L(x, \nu) = (-3 + \nu)x_1^2 + (1 + \nu)x_2^2 + (2 + \nu)x_3^2 + 2(x_1 + x_2 + x_3) - \nu.$$

The optimality conditions, as stated on page 5–22 of the lectures, are:

1. *Primal feasibility:*

$$x_1^2 + x_2^2 + x_3^2 = 1. \tag{3}$$

4. x is a minimizer of the Lagrangian. The Lagrangian is a quadratic function of x, so we can apply the results on lecture 5, page 17: x is a minimizer of $L(\cdot, \nu)$ if and only if

$$\nabla_{xx}^2 L(x, \nu) = 2 \begin{bmatrix} -3 + \nu & 0 & 0 \\ 0 & 1 + \nu & 0 \\ 0 & 0 & 2 + \nu \end{bmatrix} \succeq 0, \tag{4}$$

the vector $(2, 2, 2)$ is in the range of $\nabla_{xx}^2 L(x, \nu)$, so the linear equations

$$\nabla_x L(x, \nu) = 2 \begin{bmatrix} (-3 + \nu)x_1 + 1 \\ (1 + \nu)x_2 + 1 \\ (2 + \nu)x_3 + 1 \end{bmatrix} = 0 \tag{5}$$

are solvable, and x is any solution of this equation.

Since strong duality holds, these conditions are necessary and sufficient conditions for optimality of x, ν (see page 243). The equation (5) is solvable if $\nu \neq 2$, $\nu \neq -1$, $\nu \neq 3$:

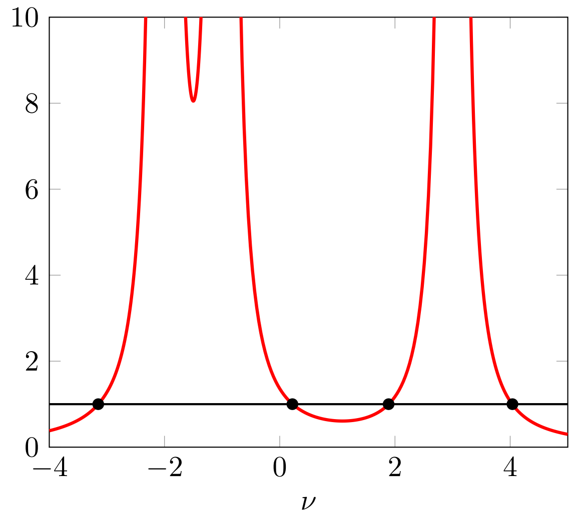
$$x_1 = \frac{-1}{\nu - 3}, \qquad x_2 = \frac{-1}{\nu + 1}, \qquad x_3 = \frac{-1}{\nu + 2}.$$

Substituting this in (3) gives a nonlinear equation in ν:

$$\frac{1}{(-3 + \nu)^2} + \frac{1}{(1 + \nu)^2} + \frac{1}{(2 + \nu)^2} = 1.$$

The left-hand side is plotted in the figure. The nonlinear equation has four solutions. To satisfy (4) we must have $\nu > 3$. The correct solution is therefore $\nu = 4.0352$. Substituting in the expression for x gives

$$x = (-0.9660, -0.1986, -0.1657).$$



6. Exercise T5.21(a,b,c).

Solution.

(a) $p^\star = 1$.

(b) The Lagrangian is $L(x, y, \lambda) = e^{-x} + \lambda x^2/y$. The dual function is

$$\begin{aligned} g(\lambda) &= \inf_{x,y>0} (e^{-x} + \lambda x^2/y) \\ &= \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0. \end{cases} \end{aligned}$$

The dual problem is

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

The dual optimal value is $d^\star = 0$.

(c) Slater's condition is not satisfied.

7. Exercise T5.17.

8. Exercise A5.3.

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

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

10. Excercise A5.26.