EE5138R: Problem Set 6
Assigned: 06/03/15
Due: 20/03/15

1. BV Problem 5.3

   Solution: For $\lambda = 0$, $g(\lambda) = \inf c^T x = -\infty$. For $\lambda > 0$,

$$\begin{aligned} g(\lambda) &= \inf\{c^T x + \lambda f(x)\} \\ &= \lambda \inf\{(c/\lambda)^T x + f(x)\} \\ &= -\lambda f_1^*(-c/\lambda) \end{aligned}$$

   The dual problem is

$$\max \ -\lambda f_1^*(-c/\lambda), \quad \text{s.t.} \quad \lambda \succeq 0.$$

2. BV Problem 5.5

   Solution: The Lagrangian is

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b) \\ &= (c^T + \lambda^T G + \nu^T A)x - \lambda^T h - \nu^T b \end{aligned}$$

   Thus the dual function is

$$g(\lambda, \nu) = \begin{cases} -\lambda^T h - \nu^T b & c^T + \lambda^T G + \nu^T A = 0 \\ -\infty & \text{else} \end{cases}$$

   The dual problem is

$$\max \ -\lambda^T h - \nu^T b \quad \text{s.t.} \quad c + G^T \lambda + A^T \nu = 0, \quad \lambda \succeq 0.$$

3. BV Problem 5.11

   Solution: The Lagrangian is

$$L(x, y_1, \ldots, y_N, z_1, \ldots, z_N) = \sum_{i=1}^{n} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 - \sum_{i=1}^{N} z_i^T (y_i - A_i x - b_i).$$

   Minimizing over y_i, we observe that

$$\inf_{y_i}\{\|y_i\|_2 + z_i^T y_i\} = \begin{cases} 0 & \|z_i\| \le 1 \\ -\infty & \text{else} \end{cases}$$

   We can minimize over x by setting the gradient with respect to x equal to zero. This yields

$$x = x_0 + \sum_{i=1}^{n} A_i^T z$$

Substituting into the Lagrangian gives

$$g(z_1,\dots,z_N) = \begin{cases} \sum_{i=1}^N (A_i x_0 + b_i)^T z_i - \frac{1}{2}\|\sum_{i=1}^N A_i^T z_i\|_2^2 & \|z_i\|_2 \le 1, i=1,\dots N \\ -\infty & \text{else} \end{cases}$$

Thus the dual problem is

$$\max \; \sum_{i=1}^N (A_i x_0 + b_i)^T z_i - \frac{1}{2}\left\|\sum_{i=1}^N A_i^T z_i\right\|_2^2 \quad \text{s.t.} \quad \|z_i\|_2 \le 1, i=1,\dots N$$

4. BV Problem 5.21

   Solutions

   (a) $p^* = 1$.

   (b) Lagrangian is

   $$L(x,y,\lambda) = e^{-x} + \lambda x^2/y.$$

   The dual function is

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

   so the duality optimal value is $d^* = \sup\{g(\lambda) : \lambda \ge 0\} = 0$.

   (c) Slater's condition is not satisfied.

   (d) $p^*(u) = 1$ if $u = 0$, $p^*(u) = 0$ if $u > 0$ and $p^*(u) = \infty$ if $u < 0$.

5. 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

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

   (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^*.

6. The relative entropy between two vectors $x, y \in \mathbf{R}^n_{++}$ is defined as

$$\sum_{k=1}^{n} x_k \log\left(\frac{x_k}{y_k}\right).$$

This is a convex function, jointly in x and y. In the following problem we calculate the vector x that minimizes the relative entropy with a given vector y, subject to equality constraints on x:

$$\min_{x \in \mathbf{R}^n} \sum_{k=1}^{n} x_k \log\left(\frac{x_k}{y_k}\right), \quad \text{s.t.} \quad Ax = b, \mathbf{1}^T x = 1$$

The domain is $\mathbf{R}^n_{++}$ and the given parameters are $y \in \mathbf{R}^n_{++}, A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$. Note that $\mathbf{1}^T x = 1$ means that x is a probability vector.

Derive the Lagrange dual of this problem and simplify it to get

$$\max_{z \in \mathbf{R}^m} \; b^T z - \log \sum_{k=1}^{n} y_k e^{a_k^T z}$$

where a_k is the k-th column of A. Note that $z \in \mathbf{R}^m$ is the Lagrange multiplier associated with the constraint $Ax = b$.

Solution: The Lagrangian is

$$L(x, \nu, z) = \sum_{k=1}^{n} x_k \log\left(\frac{x_k}{y_k}\right) - \nu(\mathbf{1}^T x - 1) - z^T(Ax - b)$$

The optimal x of the Lagrangian has k-th component given by

$$x_k \frac{1}{x_k} + \log\left(\frac{x_k}{y_k}\right) - \nu - a_k^T z = 0.$$

Thus the optimal x_k is

$$x_k^* = \frac{1}{Z} y_k e^{a_k^T z}$$

where Z is a constant in terms of ν that allows x to sum to one, i.e.,

$$Z = \sum_k y_k e^{a_k^T z}.$$

Plugging this back into the Lagrangian, we have

$$g(z) = L(x^*, \nu, z) = \sum_k \frac{1}{Z} y_k e^{a_k^T z} \log\left(\frac{e^{a_k^T z}}{Z}\right) - z^T \left(A \begin{bmatrix} \frac{1}{Z} y_1 e^{a_1^T z} \\ \vdots \\ \frac{1}{Z} y_n e^{a_n^T z} \end{bmatrix} - b \right)$$

Simplifying, we get

$$g(z) = -\log Z + b^T z$$

as desired.