HW6 Solutions

December 24, 2019

1. The Lagrangian is

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

We first minimize over y_i. We have

$$\inf_{y_i}(\|y_i\|_2 + z_i^T y_i) = \begin{cases} 0, & \|z_i\|_2 \le 1 \\ -\infty, & \text{otherwise.} \end{cases}$$

(If $\|z_i\|_2 > 1$, choose $y_i = -tz_i$ and let $t \to \infty$, to show that the function is unbounded below. If $\|z_i\|_2 \le 1$, it follows from the Cauchy- Schwarz inequality that $\|y_i\|_2 + z_i^T y_i \ge 0$, so the minimum is reached when $y_i = 0$.)

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 in the Lagrangian gives the dual function

$$g(z_1, \cdots, z_N) = \begin{cases} \sum\limits_{i=1}^{N}(A_i x_0 - b_i)^T z_i - \frac{1}{2}\left\|\sum\limits_{i=1}^{N} A_i^T z_i\right\|_2^2, & \|z_i\|_2 \le 1,\ i = 1, \ldots, N; \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & \sum\limits_{i=1}^{N}(A_i x_0 - b_i)^T z_i - \frac{1}{2}\left\|\sum\limits_{i=1}^{N} A_i^T z_i\right\|_2^2 \\ \text{subject to} & \|z_i\|_2 \le 1,\ i = 1, \ldots, N. \end{array}$$

2. (a) We show that this is a convex problem as follows

$$f_0(x) = e^{-x},\ f_0^{"} = e^{-x} \ge 0,$$

and

$$\frac{x^2}{y}\ \textit{is convex for}\ y \ge 0. \quad (\textit{p.73 of the textbook})$$

Furthermore, the optimal value and optimal point are $x^* = 0,\ p^* = 1.$

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

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

so we can write the dual problem as

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

with optimal value $d^* = 0$.

(c) The optimal duality gap is $p^* - d^* = 1$. Slater's condition is not satisfied.

3. Clearly, $x^* = (1, 1, 1, 1)$ is feasible (it satisfies the first four constraints with equality).

The Lagrangian is $L(x, z) = c^T x + z^T (Ax - b)$. The dual function is

$$g(z) = \inf(c^T x + z^T (Ax - b)) = \begin{cases} -z^T b, & z \succeq 0,\ c + A^T z = 0; \\ -\infty, & \mathit{otherwise}. \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -z^T b \\ \text{subject to} & c + A^T z = 0 \\ & z \succeq 0. \end{array}$$

The point $z^* = (3, 2, 2, 7, 0)$ is a certificate of optimality of $x = (1, 1, 1, 1)$:

- z^* is dual feasible: $z^* \succeq 0$ and $A^T z^* + c = 0$.
- z^* satisfies the complementary slackness condition:

$$z_i^*(a_i^T x - b_i) = 0, \quad i = 1, \ldots, m,$$

since the first four components of $Ax - b$ and the last component of z^* are zero.

4. (a) Follows from $\mathbf{tr}(W x x^T) = x^T W x$ and $(x x^T)_{ii} = x_i^2$

(b) It gives a lower bound because we minimize the same objective over a larger set. If X is rank one, it is optimal.

(c) We write the problem as a minimization problem

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0. \end{array}$$

Introducing a Lagrange multiplier $X \in \mathbf{S}^n$ for the matrix inequality, we obtain the Lagrangian

$$\begin{aligned} L(\nu, X) &= \mathbf{1}^T \nu - \mathbf{tr}(X(W + \mathbf{diag}(\nu))) \\ &= \mathbf{1}^T \nu - \mathbf{tr}(XW) - \sum_{i=1}^n \nu_i X_{ii} \\ &= -\mathbf{tr}(XW) + \sum_{i=1}^n \nu_i (1 - X_{ii}). \end{aligned}$$

This is bounded below as a function of ν only if $X_{ii} = 1$ for all i, so we obtain the dual problem

$$\begin{array}{ll} \text{maximize} & -\mathbf{tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \ i = 1, \ldots, n. \end{array}$$

Changing the sign again, and switching from maximization to minimization, yields the problem in part (a).

5. $f(x) = \log(e^x + e^{-x})$ is a smooth convex function, with a unique minimum at the origin. The pure Newton method started at $x^{(0)} = 1$ produces the following sequence.

k	$x^{(k)}$	$f(x^{(k)}) - p^*$
1	$-8.134\ e\{-01\}$	$2.997\ e\{-01\}$
2	$4.094\ e\{-01\}$	$8.156\ e\{-02\}$
3	$-4.730\ e\{-02\}$	$1.118\ e\{-03\}$
4	$7.060\ e\{-05\}$	$2.492\ e\{-09\}$

Started at $x^{(0)} = 1.2$, the method diverges.

k	$x^{(k)}$	$f(x^{(k)}) - p^*$
1	-1.5331	$8.8551\ e\{-01\}$
2	3.82044	3.1278
3	$-5.1658\ e\{+02\}$	$5.1588\ e\{+02\}$
4	∞	∞