Problem Set 8 Solutions

April 26, 2005

4.2.1

We consider the problem

$$\text{minimize} \;\; f(x) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2$$

$$\text{subject to } x_2 = 0.$$

(a) We have

$$L(x,\lambda) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \lambda x_2,$$

so

$$\nabla_x L(x,l) = \begin{pmatrix} x_1 \\ x_2 - 3 + \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \nabla_\lambda L(x,l) = x_2 = 0.$$

The only candidate for optimality is $(x_1^*, x_2^*) = (0,0)$ with the corresponding Lagrange multiplier $\lambda^* = 3$. Since $f(x)$ is convex over the constraint set, the point $(0,0)$ is the optimal solution.

(b) The augmented Lagrangian is

$$\begin{aligned} L_c(x,\lambda) &= \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + lx_2 + \frac{c}{2}x_2^2 \\ &= \frac{1}{2}x_1^2 + \frac{c-1}{2}x_2^2 + (l-3)x_2 \\ &= \frac{1}{2}x_1^2 + (\frac{c-1}{2}x_2 + l - 3)x_2. \end{aligned}$$

This function has a minimum for $c^k > 1$:

$$\nabla_x L_{c^k}(x^k, \lambda^k) = \begin{pmatrix} x_1^k \\ c^k - 1)x_2^k + \lambda^k - 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

so that

$$x_1^k = 0, \quad x_2^k = \frac{3 - \lambda^k}{c^k - 1} \tag{1}$$

and the corresponding optimal value $L_{c^k}(x^k, \lambda^k)$ is

$$L_{c^k}(x^k, \lambda^k) = (\frac{c^k - 1}{2} x_2^k + \lambda^k - 3)x_2^k = \frac{\lambda^k - 3}{2} \cdot \frac{3 - \lambda^k}{c^k - 1} = -\frac{1}{2} \cdot \frac{(3 - \lambda^k)^2}{c^k - 1}. \qquad (2)$$

The results for the *Quadratic penalty method* with $\lambda^k = 0$ for all k are given in the following table:

k	x_1^k	x_2^k	$L_{c^k}(x^k, \lambda^k)$
0	0	0.3333333	-0.5000000
1	0	0.0303030	-0.0454545
2	0	0.0030030	0.0045045

For the method of multipliers, the optimal point of the augmented Lagrangian $L_{c^k}(x^k, \lambda^k)$ and the optimal value of the augmented Lagrangian are still given by Eqs. (1) and (2). The only difference here is that λ^k in these equations is updated according to

$$\lambda^{k+1} = \lambda^k + c^k h(x^k) = \lambda^k + c^k x_2^k, \qquad k = 0, 1, 2,$$

where λ^0 is an initial multiplier value. The results for the *Multiplier method* with $\lambda^0 = 0$ are given in the following table:

k	λ	x_1^k	x_2^k	$L_{c^k}(x^k, \lambda^k)$
0	0	0	0.3333333	-0.5000000
1	3.3333333	0	-0.0033670	-0.0005611
2	2.9966329	0	0.0000033	-5.67$\times 10^{-9}$

By comparing the values of $L_{c^k}(x^k, \lambda^k)$, from the above results we see that the convergence of the multiplier method is significantly faster than that of the quadratic penalty method. (c) See attached plots. We have

$$p(u) = \min_{x_2 = u} \frac{1}{2}(x_1^2 + x_2^2) - 3x_2 = -\frac{1}{2}u^2 - 3u.$$

(d) For the augmented Lagrangian to have a minimum, we need $c + \nabla^2 p(0) = c - 1$ to be positive, so that $c > 1$. For the multiplier method with the constant c, we have

$$\lambda^{k+1} = \lambda^k + ch(x^k) = \frac{-\lambda^k + 3c}{c - 1}.$$

For $\{\lambda^k\}$ to converge to λ^*, we require that

$$\frac{|\lambda^{k+1} - \lambda^*|}{|\lambda^k - \lambda^*|} < 1.$$

Since $l^* = 3$, it follows that

$$\frac{|\lambda^{k+1} - \lambda^*|}{|\lambda^k - \lambda^*|} = \frac{|\frac{-\lambda^k + 3c}{c-1} - 3|}{|\lambda^k - 3|} = \frac{1}{|c - 1|},$$

and the convergence takes place when

$$\frac{1}{|c-1|} < 1.$$

Because $c > 1$, this relation reduces to $c > 2$.

5.1.1

Consider the problem

$$\text{minimize} \;\; x_1$$
$$\text{subject to } |x_1| + |x_2| \leq 1, \quad x \in X = \Re^2$$

(cf. Figure 1).

We have

$$L(x, \mu) = x_1 + \mu(|x_1| + |x_2| - 1)$$

and so

$$q(\mu) = \inf_{x \in \Re^2} L(x, \mu) = \inf_{x \in \Re^2} \{-\mu + x_1 + \mu|x_1| + \mu|x_2|\}.$$

If $0 \leq \mu < 1$, then $q(\mu)$ can be made arbitrarily small by making x_1 small. Otherwise, $q(\mu)$ is minimized by setting x_1 and x_2 to 0. Thus

$$q(\mu) = \begin{cases} -\infty & \text{if } 0 \leq \mu < 1, \\ -\mu & \text{if } 1 \leq, \end{cases}$$

and

$$q^* = \max_{\mu \geq 0} q(\mu) = -1$$

is attained at $\mu^* = 1$. Thus the only optimal solution is $x^* = (-1, 0)$ and the only Lagrange multiplier is $\mu^* = 1$. The dual function is given in Figure 2.

Consider the problem

$$\text{minimize} \;\; x_1$$
$$\text{subject to } |x_1| + |x_2| \leq 1, \quad x \in X = \{x \mid |x_1| \leq 1, \; |x_2| \leq 1\}$$

(cf. Figure 3).

We have the same Lagrangian function as before and so

$$q(\mu) = \inf_{-1 \leq x_1, x_2 \leq 1} L(x, \mu) = \begin{cases} -1 & \text{if } 0 \leq \mu \leq 1, \\ -\mu & \text{if } 1 \leq \mu. \end{cases}$$

Thus the dual optimal value is $q^* = -1$, and every $\mu^* \in [0, 1]$ is a dual optimal solution. From Figure 3, the only optimal solution is $x^* = (-1, 0)$ and corresponding Lagrange multipliers are $\mu^* \in [0, 1]$. The dual function is given in Figure 4.

5.1.2

(a) The problem is

$$\text{minimize} \;\; 10x_1 + 3x_2$$

$$\text{subject to } 5x_1 + x_2 \geq 4, \quad x_1, x_2 = 0 \text{ or } 1.$$

(b) The Lagrangian function is

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

and the dual function is

$$q(\mu) = \inf_{x_1, x_2 \in \{0,1\}} \{4\mu + (10 - 5\mu)x_1 + (3 - \mu)x_2\} = \begin{cases} 4\mu & \text{if } 0 \leq \mu \leq 2, \\ 10 - \mu & \text{if } 2 \leq \mu \leq 3, \\ 13 - 2\mu & \text{if } 3 \leq \mu, \end{cases}$$

.

(c) From (a), we see that $x^* = (1, 0)$ and $f^* = 10$. From (b), we see that $q^* = 8$. Thus there is a duality gap of $f^* - q^* = 2$ and there is no Lagrange multiplier.

5.1.3

A straightforward calculation yields the dual function as

$$q(\lambda) = \min_{x \in} \{\|z - x\|^2 + \lambda' Ax\} = -\frac{\|A'\lambda\|^2}{4} + \lambda' Az.$$

Thus the dual problem is equivalent to

$$\min_{\lambda \in^m} \{\frac{\|A'\lambda\|^2}{4} - \lambda' Az + \|z\|^2\}$$

or

$$\min_{\lambda \in^m} \|z - \frac{A'\lambda}{2}\|^2.$$

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

5.1.5

Obviously the primal LP is infeasible since we simultaneously require $x_1 \geq 0$ and $x_1 \leq -1$. We write the Lagrangian as

$$L(x, \mu) = x_1 - x_2 + \mu_1(x_1 + 1) + \mu_2(1 - x_1 - x_2) = (1 + \mu_1 - \mu_2)x_1 + (-1 - \mu_2)x_2 + \mu_1 + \mu_2.$$

Now the dual function is computed as

$$q(\mu) = \begin{cases} \mu_1 + \mu_2 & \text{if } 1 + \mu_1 - \mu_2 \geq 0, \ -1 - \mu_2 \geq 0 \\ -\infty & \text{otherwise} \end{cases}.$$

We then maximize $q(\mu)$ over all $\mu_1 \geq 0$, $\mu_2 \geq 0$, which gives the dual LP specified. Again this is clearly infeasible since we simultaneously require $\mu_2 \geq 0$ and $\mu_2 \leq -1$.