Solutions Chapter 6

SECTION 6.3

6.3.1 (www)

(a) Let μ^* be a dual optimal solution. Similar to the proof of Prop. 6.3.1, we obtain

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - 2s^k\big(q^* - q(\mu^k)\big) + (s^k)^2||g^k||^2,$$

where $q^* = q(\mu^*)$. Since $s^k = \frac{q^* - q(\mu^k)}{||g^k||^2}$, we have

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - \frac{\big(q^* - q(\mu^k)\big)^2}{||g^k||^2}. \tag{1}$$

Therefore

$$||\mu^{k+1} - \mu^*|| < ||\mu^k - \mu^*||, \qquad \forall\ k,$$

implying that $\{\mu^k\}$ is bounded.

(b) Let C be a positive constant such that $||g^k|| \leq C$ for all k. Then from Eq. (1) it follows that

$$||\mu^{k+1} - \mu^*||^2 + \frac{\big(q^* - q(\mu^k)\big)^2}{C^2} \leq ||\mu^k - \mu^*||^2, \quad \forall\ k.$$

By summing these inequalities over all k, we obtain

$$\frac{1}{C^2}\sum_{k=0}^{\infty}\big(q^* - q(\mu^k)\big)^2 \leq ||\mu^0 - \mu^*||^2,$$

so that

$$\lim_{k\to\infty} q(\mu^k) = q^*. \tag{2}$$

Since $\{\mu^k\}$ is bounded, there exist a vector $\hat{\mu}$ and a subsequence $\{\mu^k\}_{k\in\mathcal{K}} \subset \{\mu^k\}$ converging to $\hat{\mu} \in M$ (set M is closed). By using the upper-semicontinuity of q, we have

$$\limsup_{k\to\infty,\ k\in\mathcal{K}} q(\mu^k) \leq q(\hat{\mu}) \leq q^*,$$

which in view of Eq. (2) implies that $q(\hat{\mu}) = q^*$. Thus every limit point of $\{\mu^k\}$ is optimal.

Now we show that $\{\mu^k\}$ actually converges. Let M^* denote the set of all dual optimal solutions. Note that M^* is convex (by concavity of q) and closed (by upper-semicontinuity of q). Suppose that $\{\mu^k\}$ has two distinct limit points, say $\hat{\mu} \in M^*$ and $\tilde{\mu} \in M^*$. As seen in (a), for any $\mu^* \in M^*$, the sequence $\{||\mu^k - \mu^*||\}$ decreases monotonically, and therefore it converges. Hence $||\hat{\mu} - \mu^*|| = ||\tilde{\mu} - \mu^*||$ for all $\mu^* \in M^*$, implying that $\tilde{\mu} = \hat{\mu}$.

(c) Let q be real-valued and concave over the entire space $\Re^r$. According to Prop. B.24 of Appendix B, since $\{\mu^k\}$ is bounded, the set $\cup_{k\geq 0}\partial q(\mu^k)$ is bounded, and so is $\{g^k\}$.

6.3.2 (www)

(a) Let $\tilde{q}$ be an underestimate of q^* such that $q(\mu^k) < \tilde{q} \leq q^*$. Consider the function $\bar{q}(\mu) = \min\big\{q(\mu), \tilde{q}\big\}$. Note that $\bar{q}$ is concave and that $\max_{\mu\in M} \bar{q}(\mu) = \tilde{q}$. The proposed method is obtained by applying the method described in Exercise 6.3.1 to the problem $\max_{\mu\in M} \bar{q}(\mu)$. The algorithm will either stop at some iteration $\bar{k}$ for which $\bar{q}(\mu^{\bar{k}}) = \tilde{q}$ [i.e., $q(\mu^{\bar{k}}) \geq \tilde{q}$] or generate a sequence $\{\mu^k\}$ such that $\bar{q}(\mu^k) = q(\mu^k) < \tilde{q}$ for all k. According to the results of Exercise 6.3.1, the sequence $\{\mu^k\}$ is bounded. Furthermore, provided that $\{g^k\}$ is bounded, the sequence $\{\mu^k\}$ converges to some point $\bar{\mu}$ such that $\bar{q}(\bar{\mu}) = \tilde{q}$. Since $q(\bar{\mu}) \geq \bar{q}(\bar{\mu})$, we have $q(\bar{\mu}) \geq \tilde{q}$.

(b) Let $\tilde{q}$ be an overestimate of q^*, and let L be a constant such that $||g^k|| \leq L$ for all k. Then for any $N > 0$, we have

$$\begin{aligned}\sum_{k=0}^{N} s^k||g^k|| &= \sum_{k=0}^{N} \frac{\tilde{q} - q(\mu^k)}{||g^k||} \\ &\geq \frac{1}{L}\sum_{k=0}^{N}(\tilde{q} - q(\mu^k)) \\ &\geq \frac{1}{L}\sum_{k=0}^{N}(\tilde{q} - q^*) \\ &= \frac{(N+1)(\tilde{q} - q^*)}{L},\end{aligned}$$

where the last inequality follows from the fact that $q(\mu^k) \leq q^*$ for all k. Since $\tilde{q} > q^*$, by taking limit as $N \to \infty$ in the expression above, we obtain the desired result.

6.3.3 (www)

(a) To obtain a contradiction, suppose that $\liminf_{k\to\infty} \sqrt{k}\big(q^* - q(\mu^k)\big) > 0$. Then there is an $\epsilon > 0$ and large enough $\bar{k}$ such that $\sqrt{k}\big(q^* - q(\mu^k)\big) \geq \epsilon$ for all $k \geq \bar{k}$. Therefore

$$\big(q^* - q(\mu^k)\big)^2 \geq \frac{\epsilon^2}{k}, \qquad \forall\, k \geq \bar{k},$$

implying that

$$\sum_{k=\bar{k}}^{\infty} \big(q^* - q(\mu^k)\big)^2 \geq \epsilon^2 \sum_{k=\bar{k}}^{\infty} \frac{1}{k} = \infty,$$

which contradicts the relation

$$\sum_{k=0}^{\infty} \big(q^* - q(\mu^k)\big)^2 < \infty$$

shown in solution of Exercise 6.3.1.

(b) As seen in Exercise 6.3.1, we have for all dual optimal solutions μ^* and all k

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - \frac{\big(q^* - q(\mu^k)\big)^2}{||g^k||^2}.$$

This relation and the inequality $q(\mu^*) - q(\mu^k) \geq a||\mu^* - \mu^k||$ yield for all k

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - \frac{a^2||\mu^k - \mu^*||^2}{||g^k||^2},$$

from which, by using $\sup_{k\geq 0} ||g^k|| \leq b$, we obtain

$$||\mu^{k+1} - \mu^*||^2 \leq \left(1 - \frac{a^2}{b^2}\right) ||\mu^k - \mu^*||^2,$$

and the desired relation follows.

6.3.4 www

Assume that w^k is not an ascent direction. Then

$$g^{k\prime} w^k = q'(\mu; w^k) \leq 0. \tag{1}$$

Since w^k is the projection of the zero vector on the set $\text{conv}\{g^1, \ldots, g^{k-1}\}$, by Prop. B.11 of Appendix B, we have

$$(g - w^k)' w^k \geq 0, \qquad \forall\ g \in \text{conv}\{g^1, \ldots, g^{k-1}\}.$$

Therefore $g^{i\prime} w^k \geq ||w^k||^2$ for all $i = 1, \ldots, k-1$. Since w^k is a subgradient of q at μ and μ is not an optimal point, we must have $||w^k|| \geq ||g^*|| > 0$, where g^* is a subgradient of q at μ that has minimum norm. Hence

$$g^{i\prime} w^k \geq ||g^*|| > 0, \qquad \forall\ i = 1, \ldots, k-1. \tag{2}$$

Suppose that the process does not terminate in a finite number of steps. Let $\{w^k, g^k\}$ be a sequence generated by the algorithm. Since $\{w^k\} \subset \partial q(\mu)$, $\{g^k\} \subset \partial q(\mu)$ and $\partial q(\mu)$ is compact, there must exist subgradients $\hat{w}, \hat{g} \in \partial q(\mu)$ and subsequences $\{w^k\}_{k\in K}, \{g^k\}_{k\in K}$ of $\{w^k\}$ and $\{g^k\}$, respectively, such that

$$\lim_{k\to\infty, k\in K} w^k = \hat{w} \quad \text{and} \quad \lim_{k\to\infty, k\in K} g^k = \hat{g}. \tag{3}$$

From Eqs. (1) and (3) we have $\hat{g}'\hat{w} \leq 0$. On the other hand, by considering Eq. (2) for $i < k$ and $i \in K$, taking the limit as $k \to \infty$, and using Eq. (3), we obtain $\hat{g}'\hat{w} > 0$ - a contradiction. Therefore the process has to terminate in a finite number of steps.

6.3.5 (www)

(a) We have

$$g \in \partial_\epsilon q(\mu) \quad \text{if and only if} \quad q(\mu + sd) \leq q(\mu) + sg'd + \epsilon, \quad \forall\, s > 0, \quad d \in \Re^r.$$

Hence

$$g \in \partial_\epsilon q(\mu) \quad \Longleftrightarrow \quad \sup_{s>0} \frac{q(\mu + sd) - q(\mu) - \epsilon}{s} \leq g'd, \quad \forall\, d \in \Re^r. \tag{1}$$

It follows that $\partial_\epsilon q(\mu)$ is the intersection of the closed subspaces

$$\left\{ g \;\middle|\; \sup_{s>0} \frac{q(\mu + sd) - q(\mu) - \epsilon}{s} \leq g'd \right\}$$

as d ranges over $\Re^r$. Hence $\partial_\epsilon q(\mu)$ is closed and convex. To show that $\partial_\epsilon q(\mu)$ is also bounded, suppose to arrive at a contradiction that there is a sequence $\{g^k\} \subset \partial_\epsilon q(\mu)$ with $||g^k|| \to \infty$. Let $d^k = -\frac{g^k}{||g^k||}$. Then, from Eq. (1), we have

$$\sup_{s>0} \frac{q(\mu + sd) - q(\mu) - \epsilon}{s} \leq -||g^k||, \qquad \forall\, s > 0,$$

so that for $s = 1$ we obtain

$$q(\mu + d^k) \to -\infty.$$

This is a contradiction since q is concave and hence continuous, so it is bounded on any bounded set. Thus $\partial_\epsilon q(\mu)$ is bounded.

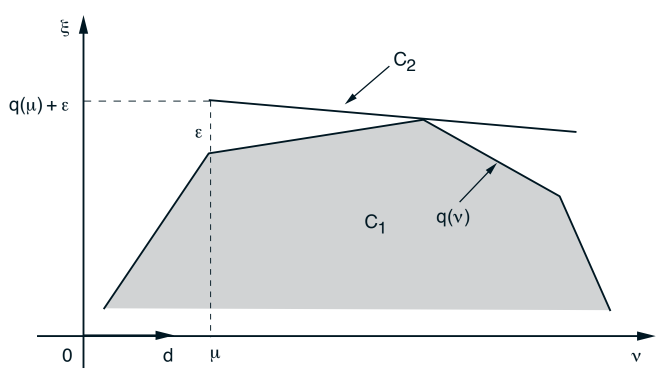

Figure for Exercise 6.3.5.

To show that $\partial_\epsilon q(\mu)$ is nonempty and satisfies

$$\sup_{s>0} \frac{q(\mu+sd)-q(\mu)-\epsilon}{s} = \inf_{g\in\partial_\epsilon q(\mu)} g'd, \qquad \forall\ d\in\Re^r,$$

we argue similar to the proof of Prop. B.24(b). Consider the subset of $\Re^{r+1}$

$$C_1 = \big\{(\xi,\nu)\mid \xi < q(\nu)\big\},$$

and the half-line

$$C_2 = \{(\xi,\nu)\mid \xi = q(\mu)+\epsilon+\alpha\sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s}, \nu=\mu+\alpha d, \alpha\ge 0\},$$

(see the figure). These sets are nonempty and convex. They are also disjoint, since we have for all $(\xi,\nu)\in C_2$

$$\xi = q(\mu)+\epsilon+\alpha\sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s} \ge q(\mu)+\epsilon+\alpha\frac{q(\mu+\alpha d)-q(\mu)-\epsilon}{\alpha} = q(\mu+\alpha d) = q(\nu).$$

Hence there exists a hyperplane separating them, i.e., a $(\gamma,w)\neq(0,0)$ such that

$$\gamma\xi+w'\nu \ge \gamma\left(q(\mu)+\epsilon+\alpha\sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s}\right)+w'(\mu+\alpha d), \quad \forall\ \alpha\ge 0,\ \ \nu\in\Re^r,\ \ \xi<q(\nu). \tag{2}$$

We cannot have $\gamma>0$, since then the left-hand side above could be made arbitrarily small by choosing ξ to be sufficiently small. Also, if $\gamma=0$, Eq. (2) implies that $w=0$, contradicting the fact that $(\gamma,w)\neq(0,0)$. Therefore, $\gamma<0$ and after dividing Eq. (2) by γ, we obtain

$$\xi+\left(\frac{w}{\gamma}\right)'(\nu-\mu) \le q(\mu)+\epsilon+\alpha\sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s}+\alpha\left(\frac{w}{\gamma}\right)'d, \quad \forall\ \alpha\ge 0,\ \ \nu\in\Re^r,\ \ \xi<q(\nu). \tag{3}$$

Taking the limit above as $\xi\to q(\nu)$ and setting $\alpha=0$, we obtain

$$q(\nu)\le q(\mu)+\epsilon+\left(-\frac{w}{\gamma}\right)'(\nu-\mu), \qquad \forall\ \nu\in\Re^r.$$

Hence $-\frac{w}{\gamma}$ belongs to $\partial_\epsilon q(\mu)$. Also by taking $\nu=\mu$ in Eq. (3), and by letting $\xi\to q(\nu)$ and by dividing with α, we obtain

$$-\frac{w}{\gamma}'d \le \frac{\epsilon}{\alpha}+\sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s},$$

Since α can be chosen as large as desired, we see that

$$-\frac{w}{\gamma}'d \le \sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s}.$$

Combining this relation with Eq. (1), we obtain

$$\min_{g\in\partial_\epsilon q(\mu)} g'd = \sup_{s>0}\frac{q(\mu+sd)-q(\mu)-\epsilon}{s}.$$

(b) By definition, $0 \in \partial_\epsilon q(\mu)$ if and only if $q(\bar{\mu}) \leq q(\mu) + \epsilon$ for all $\bar{\mu} \in \Re^r$, which is equivalent to $\sup_{\bar{\mu} \in \Re^r} q(\bar{\mu}) - \epsilon \leq q(\mu)$.

(c) Assume that a direction d is such that

$$\inf_{g \in \partial_\epsilon q(\mu)} d'g > 0, \tag{1}$$

while $\sup_{s>0} q(\mu + sd) \leq q(\mu) + \epsilon$. Then $q(\mu + sd) - q(\mu) \leq \epsilon$ for all $s > 0$, or equivalently

$$\frac{q(\mu + sd) - q(\mu) - \epsilon}{s} \leq 0, \qquad \forall\ s > 0.$$

Consequently, using part (a), we have

$$\inf_{g \in \partial_\epsilon q(\mu)} d'g = \sup_{s>0} \frac{q(\mu + sd) - q(\mu) - \epsilon}{s} \leq 0.$$

which contradicts Eq. (1).

(d) The vector $g_\mu \in \partial_\epsilon q(\mu)$ such that $||g_\mu|| = \min_{g \in \partial_\epsilon q(\mu)} ||g||$ is the projection of the zero vector on the convex and compact set $\partial_\epsilon q(\mu)$. If $0 \notin \partial_\epsilon q(\mu)$, we have $\|g_\mu\| > 0$. By Prop. B.11 of Appendix B, we have

$$(g - g_\mu)'g_\mu \geq 0, \qquad \forall\ g \in \partial_\epsilon q(\mu).$$

Hence

$$g'g_\mu \geq ||g_\mu||^2 > 0, \qquad \forall\ g \in \partial_\epsilon q(\mu).$$

(e) Let g_1 be some ϵ-subgradient of q. For $k = 2, 3, \ldots$, let w^k be the vector of minimum norm in the convex hull of $g^1, \ldots, g^{k-1}$,

$$w^k = \arg \min_{g \in conv\{g^1, \ldots, g^{k-1}\}} \|g\|.$$

If $w^k = 0$, stop; we have $0 \in \partial_\epsilon q(\mu)$, so $\sup_{\bar{\mu} \in \Re^r} q(\bar{\mu}) - \epsilon \leq q(\mu)$. Otherwise, by search along the line $\{\mu + sw^k \mid s \geq 0\}$, determine whether there exists a scalar $\bar{s}$ such that $q(\mu + \bar{s}w^k) > q(\mu) - \epsilon$. If such a $\bar{s}$ can be found, stop and go to the next iteration with μ replaced by $\mu + \bar{s}w^k$ (the dual value has been improved by at least ϵ). Otherwise let g^k be an element of $\partial_\epsilon q(\mu)$ such that

$${g^k}'w^k = \min_{g \in \partial_\epsilon q(\mu)} g'w^k.$$

From part (a), we see that ${g^k}'w^k \leq 0$. By replicating the proof of Exercise 6.3.4, we see that this process will terminate in a finite number of steps with either an improvement of the dual value by at least ϵ, or by confirmation that $\sup_{\bar{\mu} \in \Re^r} q(\bar{\mu}) - \epsilon \leq q(\mu)$, so that μ is an ϵ-optimal solution.

(f) For the case of constrained maximization redefine the ϵ-subdifferential at a vector $\mu \in M$ to be the set

$$\{g \mid q(\nu) \leq q(\mu) + g'(\nu - \mu) + \epsilon, \ \forall \ \nu \in M\}.$$

Generalization of part (a): $0 \in \partial_\epsilon q(\mu)$ if and only if $\sup_{\nu \in M} q(\nu) - \epsilon \leq q(\mu)$.

Generalization of part (b): If a feasible direction d at μ is such that $\inf_{g \in \partial_\epsilon q(\mu)} d'g > 0$, then

$$\sup_{s>0, \mu+sd \in M} q(\mu + sd) > q(\mu) + \epsilon.$$

Parts (c) and (d) are unchanged.

The proofs are similar to those given above.

6.3.6 (www)

Let μ^* be an optimal point and $q^* = q(\mu^*)$ be the optimal value. By induction, we will show that

$$(\mu^* - \mu^k)'d^k \geq (\mu^* - \mu^k)'g^k, \qquad \forall \ k. \tag{1}$$

We have $d^0 = g^0$ and Eq. (1) holds for $k = 0$. Assuming Eq. (1) holds for k, we will prove it for $k+1$. Since $d^{k+1} = g^{k+1} + \beta^k d^k$, we have

$$\begin{aligned}(\mu^* - \mu^{k+1})'d^{k+1} &= (\mu^* - \mu^{k+1})'g^{k+1} + \beta^k(\mu^* - \mu^{k+1})'d^k \\ &= (\mu^* - \mu^{k+1})'g^{k+1} + \beta^k(\mu^* - \mu^k)'d^k + \beta^k(\mu^k - \mu^{k+1})'d^k.\end{aligned} \tag{2}$$

By the concavity of q, the fact $g^k \in \partial q(\mu^k)$, and the stepsize definition $s^k \leq \frac{q^* - q(\mu^k)}{||d^k||^2}$, we obtain

$$(\mu^* - \mu^k)'g^k \geq q^* - q(\mu^k) \geq s^k||d^k||^2. \tag{3}$$

The inductive hypothesis together with Eq. (3) implies that

$$(\mu^* - \mu^k)'d^k \geq s^k||d^k||^2.$$

Substituting this estimate in Eq. (2) yields

$$\begin{aligned}(\mu^* - \mu^{k+1})'d^{k+1} &\geq (\mu^* - \mu^{k+1})'g^{k+1} + \beta^k s^k||d^k||^2 + \beta^k(\mu^k - \mu^{k+1})'d^k \\ &= (\mu^* - \mu^{k+1})'g^{k+1} + \beta^k(s^k d^k + \mu^k - \mu^{k+1})'d^k \\ &\geq (\mu^* - \mu^{k+1})'g^{k+1},\end{aligned}$$

where the last inequality follows from the properties of projection and the definition of the method. Hence Eq. (1) holds for all k.

Next we show that

$$||\mu^{k+1} - \mu^*|| < ||\mu^k - \mu^*||, \qquad \forall\ k. \tag{4}$$

Note that Eq. (3) implies that

$$(s^k)^2||d^k||^2 \le s^k(\mu^* - \mu^k)'g^k < 2s^k(\mu^* - \mu^k)'g^k. \tag{5}$$

We have

$$\begin{aligned}
||\mu^{k+1} - \mu^*||^2 &\le ||\mu^k + s^k d^k - \mu^*||^2 \\
&\le ||\mu^k - \mu^*||^2 - 2s^k(\mu^* - \mu^k)'d^k + (s^k)^2||d^k||^2 \\
&< ||\mu^k - \mu^*||^2 - 2s^k(\mu^* - \mu^k)'d^k + 2s^k(\mu^* - \mu^k)'g^k \\
&= ||\mu^k - \mu^*||^2 - 2s^k(\mu^* - \mu^k)'(d^k - g^k) \\
&\le ||\mu^k - \mu^*||^2,
\end{aligned}$$

where the first inequality follows from the nonexpansiveness property of the projection, and the second and the last inequalities follow from Eqs. (5) and (1), respectively. Thus Eq. (4) holds.

From the definitions of d^k and β^k we have

$$\begin{aligned}
||d^k||^2 - ||g^k||^2 &= ||g^k + \beta^k d^{k-1}||^2 - ||g^k||^2 \\
&= (\beta^k)^2||d^{k-1}||^2 + 2\beta^k g^{k\prime} d^{k-1} \\
&= \beta^k||d^{k-1}||^2\left(\beta^k + 2\frac{g^{k\prime} d^{k-1}}{||d^{k-1}||^2}\right) \\
&= (2-\gamma)\beta^k g^{k\prime} d^{k-1} \\
&\le 0,
\end{aligned}$$

and therefore $||d^k|| \le ||g^k||$, which combined with Eq. (1) implies that

$$\frac{(\mu^* - \mu^k)'d^k}{||d^k||} \ge \frac{(\mu^* - \mu^k)'g^k}{||g^k||}, \qquad \forall\ k.$$

6.3.12 (www)

As in the proof of Prop. 6.3.1, we have

$$||\mu^{k+1} - \mu||^2 \le ||\mu^k - \mu||^2 - 2s^k g^{k\prime}(\mu - \mu^k) + (s^k)^2||g^k||^2, \qquad \forall\ \mu \in M,$$

where $s^k = \frac{q(\mu^*) - q(\mu^k)}{||g^k||^2}$ and $g^k \in \partial_\epsilon q(\mu^k)$. From this relation and the definition of the ϵ-subgradient we obtain

$$||\mu^{k+1} - \mu||^2 \le ||\mu^k - \mu||^2 - 2s^k\big(q(\mu) - q(\mu^k) - \epsilon\big) + (s^k)^2||g^k||^2, \qquad \forall\ \mu \in M.$$

Let μ^* be an optimal solution. Substituting the expression for s^k and taking $\mu = \mu^*$ in the above inequality, we have

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - \frac{q(\mu^*) - q(\mu^k)}{||g^k||^2}\big(q(\mu^*) - q(\mu^k) - 2\epsilon\big).$$

Thus, if $q(\mu^*) - q(\mu^k) - 2\epsilon > 0$, we obtain

$$||\mu^{k+1} - \mu^*|| \leq ||\mu^k - \mu^*||.$$

6.3.13 (www)

(a) Similar to Exercise 6.3.12, we can show that

$$||\mu^{k+1} - \mu||^2 \leq ||\mu^k - \mu||^2 - 2s^k\big(q(\mu) - q(\mu^k) - \epsilon^k\big) + (s^k)^2||g^k||^2, \qquad \forall\ \mu \in M, \quad \forall\ k.$$

By rearranging terms, we can rewrite the above inequality as

$$2s^k\left(q(\mu) - q(\mu^k) - \epsilon^k - \frac{1}{2}s^k||g^k||^2\right) + ||\mu^{k+1} - \mu||^2 \leq ||\mu^k - \mu||^2, \qquad \forall\ \mu \in M, \quad \forall\ k. \quad (1)$$

Suppose that $\limsup_{k\to\infty} q(\mu^k) < \sup_{\mu\in M} q(\mu) - \epsilon$, i.e., there is a scalar $\delta > 0$ and a nonnegative integer k_0 such that

$$q(\mu^k) \leq \sup_{\mu\in M} q(\mu) - \epsilon - \delta, \qquad \forall\ k \geq k_0.$$

Choose a point $\bar{\mu} \in M$ such that

$$q(\mu^k) \leq q(\bar{\mu}) - \epsilon - \delta, \qquad \forall\ k \geq k_0. \quad (2)$$

By setting $\mu = \bar{\mu}$ in Eq. (1) and combining it with Eq. (2), we obtain

$$2s^k\left(\epsilon + \delta - \epsilon^k - \frac{1}{2}s^k||g^k||^2\right) + ||\mu^{k+1} - \bar{\mu}||^2 \leq ||\mu^k - \bar{\mu}||^2, \qquad \forall\ k \geq k_0.$$

Since $\epsilon^k \to \epsilon$ and $s^k||g^k||^2 \to 0$, we can assume that k_0 is large enough so that

$$\epsilon + \delta - \epsilon^k - \frac{1}{2}s^k||g^k||^2 \geq \frac{\delta}{2}, \qquad \forall\ k \geq k_0.$$

Therefore we have

$$s^k\delta + ||\mu^{k+1} - \bar{\mu}||^2 \leq ||\mu^k - \bar{\mu}||^2, \qquad \forall\ k \geq k_0.$$

Summation of the above inequalities gives

$$\delta\sum_{k=k_0}^{N} s^k + ||\mu^{N+1} - \bar{\mu}||^2 \leq ||\mu^{k_0} - \bar{\mu}||^2.$$

Letting $N \to \infty$ in the relation above yields $\sum_{k=k_0}^{\infty} s^k < \infty$, which is a contradiction. Therefore we must have that

$$\limsup_{k\to\infty} q(\mu^k) \geq \sup_{\mu\in M} q(\mu) - \epsilon. \tag{3}$$

On the other hand, since $q(\mu^k) \leq \sup_{\mu\in M} q(\mu)$ for all k, we have

$$\limsup_{k\to\infty} q(\mu^k) \leq \sup_{\mu\in M} q(\mu),$$

which combined with Eq. (3) yields the desired result.

(b) By rearranging the terms and setting $\epsilon^k = 0$ in Eq. (1), we obtain

$$2s^k\big(q(\mu) - q(\mu^k)\big) + ||\mu^{k+1} - \mu||^2 \leq ||\mu^k - \mu||^2 + (s^k)^2||g^k||^2, \qquad \forall\ \mu \in M,\ k. \tag{4}$$

Assume, to arrive at a contradiction, that $\limsup_{k\to\infty} q(\mu^k) < \sup_{\mu\in M} q(\mu)$. Then, by the same reasoning as in part (a), it can be seen that there is a point $\bar{\mu} \in M$ such that Eq. (2) is valid with $\epsilon = 0$. This together with Eq. (4), where $\mu = \bar{\mu}$, implies that

$$2s^k\delta + ||\mu^{k+1} - \bar{\mu}||^2 \leq ||\mu^k - \bar{\mu}||^2 + (s^k)^2||g^k||^2, \qquad \forall\ k \geq k_0.$$

Summation of these inequalities over k for $k_0 \leq k \leq N$ gives

$$2\delta \sum_{k=k_0}^{N} s^k + ||\mu^{N+1} - \bar{\mu}||^2 \leq ||\mu^{k_0} - \bar{\mu}||^2 + \sum_{k=k_0}^{N} (s^k)^2||g^k||^2, \qquad \forall\ N \geq k_0.$$

Therefore

$$2\delta \sum_{k=k_0}^{N} s^k \leq ||\mu^{k_0} - \bar{\mu}||^2 + \sum_{k=k_0}^{\infty} (s^k)^2||g^k||^2 < \infty, \qquad \forall\ N \geq k_0.$$

By letting $N \to \infty$ in the above relation, we obtain $\sum_{k=k_0}^{\infty} s^k < \infty$, which is a contradiction. Hence

$$\limsup_{k\to\infty} q(\mu^k) = \sup_{\mu\in M} q(\mu).$$

Let μ^* be an optimal point. By setting $\mu = \mu^*$ in Eq. (4) and summing the obtained inequalities for $n \leq k \leq N$, we have

$$2\sum_{k=n}^{N} s^k\big(q(\mu^*) - q(\mu^k)\big) + ||\mu^{N+1} - \mu^*||^2 \leq ||\mu^n - \mu^*||^2 + \sum_{k=n}^{N}(s^k)^2||g^k||^2, \qquad \forall\ n \leq k \leq N,$$

where $n < N$ are some positive integers. Let n be fixed and let $N \to \infty$ in the last inequality. Then

$$\limsup_{N\to\infty} ||\mu^{N+1} - \mu^*||^2 \leq ||\mu^n - \mu^*||^2 + \sum_{k=n}^{\infty}(s^k)^2||g^k||^2, \qquad \forall\ n \geq 0. \tag{5}$$

Hence $\{\mu^k\}$ is bounded. Let $\{\mu^{k_j}\} \subset \{\mu^k\}$ be a subsequence such that

$$\lim_{j\to\infty} q(\mu^{k_j}) = \limsup_{k\to\infty} q(\mu^k) = \sup_{\mu\in M} q(\mu) = q(\mu^*).$$

Without loss of generality, we may assume that $\{\mu^{k_j}\}$ converges to some point $\hat{\mu}$. The set M is closed, so that $\hat{\mu} \in M$. By the upper semicontinuity of q, we have

$$\limsup_{j\to\infty} q(\mu^{k_j}) \leq q(\hat{\mu}).$$

Since $\limsup_{j\to\infty} q(\mu^{k_j}) = \lim_{j\to\infty} q(\mu^{k_j}) = q(\mu^*)$, the relation above implies $q(\hat{\mu}) = q(\mu^*)$. Thus $\hat{\mu}$ is optimal. Then Eq. (5) is valid with $\mu^* = \hat{\mu}$ and $n = k_j$ for some j, i.e.,

$$\limsup_{N\to\infty} ||\mu^{N+1} - \hat{\mu}||^2 \leq ||\mu^{k_j} - \hat{\mu}||^2 + \sum_{k=k_j}^{\infty} (s^k)^2 ||g^k||^2.$$

By letting $j \to \infty$ and taking into account that $\lim_{j\to\infty}\left(||\mu^{k_j} - \hat{\mu}||^2 + \sum_{k=n_j}^{\infty}(s^k)^2||g^k||^2\right) = 0$ (since the boundedness of $\{\mu^k\}$ implies boundedness of $\{g^k\}$), we obtain

$$\limsup_{N\to\infty} ||\mu^{N+1} - \hat{\mu}|| = 0,$$

which implies that $\mu^k \to \hat{\mu}$.

6.3.14 (www)

(a) Let μ^* be an optimal point and let $\epsilon > 0$ be given. If $g^{\bar{k}} = 0$ for some $\bar{k}$, then $q(\mu^{\bar{k}}) = q(\mu^*)$ and one may take $\bar{\mu} = \mu^*$. If $g^k \neq 0$ for all k, then by the nonexpansiveness of the projection operation, we have

$$||\mu^{k+1} - \mu^*||^2 \leq \left\|\mu^k + \frac{\alpha g^k}{||g^k||} - \mu^*\right\|^2 = ||\mu^k - \mu^*||^2 + \alpha^2 - 2\alpha(\mu^* - \mu^k)'\frac{g^k}{||g^k||}, \qquad \forall\ k. \qquad (1)$$

Note that the term $(\mu^* - \mu^k)'g^k/||g^k||$ represents the distance from μ^* to the supporting hyperplane $H_k = \{\mu \mid {g^k}'(\mu^k - \mu) = 0\}$. Define $L_k = \{\mu \in M \mid q(\mu) = q(\mu^k)\}$. Since q is concave and real valued over the entire space, it is continuous over $\Re^r$. Therefore L_k is closed, and the distance

$$\rho^k = \min_{\mu\in L_k} ||\mu - \mu^*||$$

from μ^* to L_k is well defined. Also the set L_k and the vector μ^* lie on the same side of the hyperplane H_k. Hence every line segment joining μ^* with a point of H_k passes through L_k, and therefore

$$(\mu^* - \mu^k)'\frac{g^k}{||g^k||} \geq \rho^k, \qquad \forall\ k.$$

Using this inequality in Eq. (1), we obtain

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^k - \mu^*||^2 + \alpha^2 - 2\alpha\rho^k, \qquad \forall\ k.$$

In order to arrive at a contradiction, suppose that $\rho^k \ge \frac{\alpha}{2}(1+\epsilon)$ for all k. Then

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^k - \mu^*||^2 - \epsilon\alpha^2, \qquad \forall\ k.$$

By summing these inequalities, we have

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^0 - \mu^*||^2 - \epsilon(k+1)\alpha^2, \qquad \forall\ \ k,$$

and by letting $k \to \infty$, we obtain a contradiction. Therefore there must exist a $\bar{k}$ and a $\bar{\mu} \in L_{\bar{k}}$ such that

$$\frac{\alpha(1+\epsilon)}{2} > \rho^{\bar{k}} = \min_{\mu \in L_{\bar{k}}} ||\mu - \mu^*|| = ||\bar{\mu} - \mu^*||,$$

as desired.

(b) Let $\mu^* \in M^*$. Assume that $\mu^k \notin M^*$ for all k. Then similar to the proof in (a) we have

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^k - \mu^*||^2 + (\alpha^k)^2 - 2\alpha^k r^k, \qquad \forall\ k. \tag{2}$$

Let $a > 0$ be fixed and let $q^* = q(\mu^*)$. Consider the set $\{\mu \in M \mid q(\mu) \ge q^* - a\}$ and its boundary Γ_{q^*-a}. By assumption, the set M^* is compact, so that Γ_{q^*-a} is compact. Furthermore $M^* \cap \Gamma_{q^*-a} = \emptyset$. Hence we have

$$\rho(a) = \min_{\mu \in \Gamma_{q^*-a},\, \nu \in M^*} ||\mu - \nu|| > 0.$$

Since $\alpha^k \to 0$, one can find $N_{\rho(a)}$ such that $\alpha^k < \rho(a)$ for all $k > N_{\rho(a)}$. If $q(\mu^k) < q^* - a$, then $\rho^k > \rho(a)$ and from Eq. (1) we have

$$||\mu^{k+1} - \mu^*||^2 < ||\mu^k - \mu^*||^2 - \rho(a)\alpha^k, \qquad \forall\ k > N_{\rho(a)}. \tag{3}$$

Since $\sum_{k=0}^{\infty} \alpha^k = \infty$, there must exist $N_a > N_{\rho(a)}$ such that $q(\mu^{N_a}) \ge q^* - a$. Define

$$d(a) = \max_{\nu \in \Gamma_{q^*-a}} \min_{\mu \in M^*} ||\nu - \mu||.$$

Let $k > N_a$. If $q(\mu^{N_a}) \ge q^* - a$, then $\min_{\mu^* \in M^*} ||\mu^k - \mu^*|| \le d(a)$ and since

$$||\mu^{k+1} - \mu^*|| \le ||\mu^k + \alpha^k g^k/||g^k|| - \mu^*|| \le ||\mu^k - \mu^*|| + \alpha^k,$$

we obtain

$$\min_{\mu^* \in M^*} ||\mu^{k+1} - \mu^*|| \le d(a) + \alpha^k. \tag{4}$$

On the other hand, if $q(\mu^k) < q^* - a$ then from Eq. (3) we have

$$\min_{\mu^* \in M^*} ||\mu^{k+1} - \mu^*|| \leq \min_{\mu^* \in M^*} ||\mu^k - \mu^*||. \tag{5}$$

Combining Eqs. (4) and (5), we obtain

$$\min_{\mu^* \in M^*} ||\mu^k - \mu^*|| \leq d(a) + \max_{k > N_a} \alpha^k, \qquad \forall\ k > N_a.$$

Since $d(a) \to 0$ as $a \to 0$, for any $\delta > 0$ there exists a_δ such that $d(a_\delta) \leq \delta/2$. Also, one can find an index N_δ such that $q(\mu^{N_\delta}) \geq q^* - a_\delta$ and $\alpha^k \leq \delta/2$ for all $k > N_\delta$. Therefore

$$\min_{\mu^* \in M^*} ||\mu^k - \mu^*|| \leq \delta, \qquad \forall\ k > N_\delta,$$

showing that

$$\lim_{k \to \infty} \min_{\mu^* \in M^*} ||\mu^k - \mu^*|| = 0.$$

By continuity of q, we have $\lim_{k \to \infty} q(\mu^k) = q^*$, which completes the proof.

6.3.15 (www)

Assumption (i) guarantees that the function $q(\mu) = \sum_{i=1}^m q_i(\mu)$ is concave and has bounded level sets. Thus, the level sets of q are compact, and hence the optimal solution set M^* is nonempty and compact.

The proof of part (a) is tricky and is based on the assumptions (i) and (ii). The proof of part (b) combines the ideas of incremental gradient method analysis of Section 1.5, together with the line of proof of Exercise 6.3.14.

(a) Let $\mu^* \in M^*$ be an arbitrary optimal solution. By the nonexpansiveness property of the projection, we have

$$||\psi^{i,k} - \mu^*||^2 \leq ||\psi^{i,k-1} + \alpha^k g^{i,k} - \mu^*||^2 \leq ||\psi^{i-1,k} - \mu^*||^2 - 2\alpha^k g^{i,k\prime}(\mu^* - \psi^{i-1,k}) + (\alpha^k)^2 C^2,\ \ \forall\ i, k.$$

Since $g^{i,k\prime}(\mu^* - \psi^{i-1,k}) \geq q_i(\mu^*) - q_i(\psi^{i-1,k})$ for each i, we obtain

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - 2\alpha^k \sum_{i=1}^m \big(q_i(\mu^*) - q_i(\psi^{i-1,k})\big) + m(\alpha^k)^2 C^2, \tag{1}$$

for all $k \geq 0$ and $\mu^* \in M^*$. Let $\hat{\alpha}$ be an upper bound for α^k, and let i_0 be an index such that the level sets of q_{i_0} are bounded. Define

$$q^* = \max_{\mu \in M} q(\mu), \qquad a = \frac{m}{2}\hat{\alpha}C^2 + \sum_{i=1}^m q_i^* - q^* > 0,$$

and

$$L(a, \mu^*) = \{\nu \in M \mid q_{i_0}(\nu) \geq q_{i_0}(\mu^*) - a\}.$$

Under assumption (i) the level set $L(a, \mu^*)$ is nonempty and compact for any $\mu^* \in M^*$. Note that for any k either $q_{i_0}(\psi^{i_0-1,k}) < q_{i_0}(\mu^*) - a$ or $q_{i_0}(\psi^{i_0-1,k}) \geq q_{i_0}(\mu^*) - a$. Suppose that the former is the case. Since $q_i(\psi^{i-1,k}) \leq q_i^*$, we have

$$\begin{aligned}
\sum_{i=1}^{m}\big(q_i(\mu^*) - q_i(\psi^{i-1,k})\big) &> \sum_{i \neq i_0}\big(q_i(\mu^*) - q_i^*\big) + a \\
&= q_{i_0}^* - q_{i_0}(\mu^*) + \frac{m}{2}\hat{\alpha}C^2 \\
&\geq \frac{m}{2}\hat{\alpha}C^2.
\end{aligned}$$

By combining this relation with Eq. (1), we obtain

$$||\mu^{k+1} - \mu^*||^2 < ||\mu^k - \mu^*||^2 - \alpha^k mC^2(\hat{\alpha} - \alpha^k) \leq ||\mu^k - \mu^*||^2,$$

where the last inequality follows from $0 < \alpha^k \leq \hat{\alpha}$. Therefore

$$||\mu^{k+1} - \mu^*|| < ||\mu^k - \mu^*|| \quad \text{whenever} \quad q_{i_0}(\psi^{i_0-1,k}) < q_{i_0}(\mu^*) - a. \tag{2}$$

If $q_{i_0}(\psi^{i_0-1,k}) \geq q_{i_0}(\mu^*) - a$, then the subiterate $\psi^{i_0-1,k}$ belongs to the level set $L(a, \mu^*)$. Therefore $||\psi^{i_0-1,k} - \mu^*|| \leq diam\big(L(a, \mu^*)\big)$, where $diam(\cdot)$ denotes the diameter of a set. Since the subgradients $g^{i,k}$ are bounded, it follows that

$$||\mu^{k+1} - \mu^*|| \leq ||\mu^{k+1} - \psi^{i_0-1,k}|| + ||\psi^{i_0-1,k} - \mu^*|| \leq \hat{\alpha}mC + diam\big(L(a, \mu^*)\big).$$

Thus

$$||\mu^{k+1} - \mu^*|| \leq \hat{\alpha}mC + diam\big(L(a, \mu^*)\big) \quad \text{whenever} \quad q_{i_0}(\psi^{i_0-1,k}) \geq q_{i_0}(\mu^*) - a. \tag{3}$$

From Eqs. (2) and (3), we have

$$||\mu^k - \mu^*|| \leq \max\big\{\hat{\alpha}mC + diam\big(L(a, \mu^*)\big),\ ||\mu^0 - \mu^*||\big\} \qquad \forall\ k \geq 0,$$

which completes the proof.

(b) Here we argue similar to the proof of Exercise 6.3.14(b). Since the stepsize is bounded, the sequence of the iterates $\{\mu^k\}$ is also bounded as seen in part (a). Let

$$C_i = \max\left\{C,\ \max_{k \geq 0}\big\{||g|| \mid g \in \partial q_i(\mu^k)\big\}\right\}.$$

Note that

$$||\psi^{i,k} - \mu^k|| \le \alpha^k \sum_{j=1}^{i} C_j, \quad \forall\ i, k. \tag{4}$$

From Eq. (1) we have

$$\begin{aligned}
||\mu^{k+1} - \mu^*||^2 &\le ||\mu^k - \mu^*||^2 - 2\alpha^k \left(q(\mu^*) - q(\mu^k) + \sum_{i=1}^{m} \big(q_i(\mu^k) - q_i(\psi^{i-1,k})\big) \right) + m(\alpha^k)^2 C^2 \\
&\le ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q(\mu^*) - q(\mu^k)\big) + 2\alpha^k \sum_{i=2}^{m} C_i ||\psi^{i-1,k} - \mu^k|| + m(\alpha^k)^2 C^2 \\
&\le ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q(\mu^*) - q(\mu^k)\big) + (\alpha^k)^2 \left(2 \sum_{i=2}^{m} C_i \sum_{j=1}^{i-1} C_j + \sum_{i=1}^{m} C_i^2 \right) \\
&= ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q(\mu^*) - q(\mu^k)\big) + (\alpha^k)^2 \left(\sum_{i=1}^{m} C_i^2 \right)^2,
\end{aligned}$$

where the next-to-last inequality follows from Eq. (4), and we are using the facts $C \le C_i$ and $q_i(\psi^{i-1,k}) - q_i(\mu^k) \le g_i'(\psi^{i-1,k} - \mu^k)$ for all $g_i \in \partial q_i(\mu^k)$. Therefore

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q(\mu^*) - q(\mu^k)\big) + (\alpha^k)^2 \bar{C}^2, \quad \forall\ \mu^* \in M^*, \quad \forall\ k \ge 0, \tag{5}$$

where $\bar{C} = \sum_{i=1}^{m} C_i$. Let $a > 0$ and k_0 such that $\alpha^k \le a/\bar{C}^2$ for all $k \ge k_0$. If $q(\mu^k) < q(\mu^*) - a$ for some $k \ge k_0$, then from Eq. (5) we have

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu_k - \mu^*||^2 - \alpha^k (2a - \alpha^k \bar{C}^2),$$

and therefore

$$\big(dist(\mu^{k+1}, M^*)\big)^2 \le \big(dist(\mu^k, M^*)\big)^2 - a\alpha^k. \tag{6}$$

Note that this relation cannot hold for all $k \ge k_0$, for otherwise the condition $\sum_{k=0}^{\infty} \alpha^k = \infty$ will be violated. Hence, there is an integer $k_1 \ge k_0$ for which $q(\mu^{k_1}) \ge q(\mu^*) - a$. This means that the point μ^{k_1} belongs to the level set $L_a = \{\mu \in M \mid q(\mu) \ge q(\mu^*) - a\}$, which is compact, so that

$$dist(\mu^{k_1}, M^*) \le \max_{\mu \in L_a} dist(\mu, M^*) < \infty.$$

Denote

$$d(a) = \max_{\mu \in L_a} dist(\mu, M^*).$$

Since $||\mu^{k_1+1} - \mu^*|| \le ||\mu^{k_1} - \mu^*|| + \alpha^{k_1}\bar{C}$, we have that $dist(\mu^{k_1+1}, M^*) \le d(a) + \alpha^{k_1}\bar{C}$. Hence for $k \ge k_1$ we have

$$dist(\mu^{k+1}, M^*) < d(\mu^k, M^*) \quad \text{if} \quad q(\mu^k) < q^* - a,$$

[cf. Eq. (6)] and

$$dist(\mu^{k+1}, M^*) \le d(a) + \alpha^k \bar{C} \quad \text{if} \quad q(\mu^k) \ge q^* - a.$$

Combining these relations, we obtain

$$dist(\mu^k, M^*) \le d(a) + \bar{C} \max_{k \ge k_1} \alpha^k, \qquad \forall\ k \ge k_1.$$

Note, using Eq. (6), that the set of indices $\{k \mid q(\mu^k) \ge q(\mu^*) - a\}$ is unbounded for any choice of $a > 0$. Since $\lim_{a \to 0} d(a) = 0$, given any $\epsilon > 0$, there is $\delta > 0$ such that for $0 < a < \delta$ we have $d(a) \le \epsilon/2$. Let the index k_δ be such that $q(\mu^{k_\delta}) \ge q(\mu^*) - a$ and $\alpha^k \le \epsilon/(2\bar{C})$ for all $k \ge k_\delta$. Then $dist(\mu^k, M^*) \le \epsilon$ for all $k \ge k_\delta$, i.e. $\lim_{k\to\infty} dist(\mu^k, M^*) = 0$. The continuity of q implies that

$$\lim_{k\to\infty} q(\mu^k) = q(\mu^*) = \max_{\mu \in M} q(\mu).$$

(c) By dropping the term $2\alpha^k(q(\mu^*) - q(\mu^k))$ in Eq. (5) and by summing the obtained inequalities over k for $n \le k \le N$, we have

$$||\mu^{N+1} - \mu^*||^2 \le ||\mu^n - \mu^*||^2 + \bar{C}^2 \sum_{k=n}^{N} (\alpha^k)^2, \qquad \forall\ \mu^* \in M^*, \ \ \forall\ n, N,\ n < N. \tag{7}$$

Since $\{\mu^k\}$ is bounded, there exist $\hat{\mu}$ and $\{\mu^{k_j}\} \subset \{\mu^k\}$ such that $\lim_{j\to\infty} \mu^{k_j} = \hat{\mu}$. The set M is closed, so that $\hat{\mu} \in M$. As seen in part (b), we have $\lim_{k\to\infty} q(\mu^k) = q^*$, and therefore $\lim_{j\to\infty} q(\mu^{k_j}) = q^*$. Hence $\hat{\mu} \in M^*$. By setting $\mu^* = \hat{\mu}$ and $n = k_j$ in Eq. (7), where j is arbitrary, we obtain

$$||\mu^{N+1} - \hat{\mu}||^2 \le ||\mu^{k_j} - \hat{\mu}||^2 + \bar{C}^2 \sum_{k=k_j}^{N} (\alpha^k)^2, \qquad \forall\ N > k_j.$$

By letting first $N \to \infty$ and then $j \to \infty$, we have

$$\limsup_{N\to\infty} ||\mu^{N+1} - \hat{\mu}||^2 \le \lim_{j\to\infty} \left(||\mu^{k_j} - \hat{\mu}||^2 + \bar{C}^2 \sum_{k=k_j}^{\infty} (\alpha^k)^2 \right) = 0,$$

and therefore $\lim_{k\to\infty} ||\mu^k - \hat{\mu}|| = 0$.

6.3.16 (www)

The proof combines the arguments of the proofs of Exercise 6.3.1 and 6.3.15(b). Similar to the proof of Exercise 6.3.15(a), we have for any $\mu^* \in M^*$

$$||\mu^{k+1} - \mu^*||^2 \le ||\mu^k - \mu^*||^2 - 2\alpha^k \sum_{i=1}^{m} \big(q_i(\mu^*) - q_i(\psi^{i-1,k})\big) + m(\alpha^k)^2 C_i^2, \quad \forall\ \mu^* \in M^*, \ \ \forall\ k. \tag{1}$$

Note also that

$$||\psi^{i,k} - \mu^k|| \leq \alpha^k \sum_{j=1}^{i} C_j, \quad \forall\ i, k. \tag{2}$$

From Eq. (1) we have [as in the proof of Exercise 6.3.15(b)]

$$\begin{aligned} ||\mu^{k+1} - \mu^*||^2 &\leq ||\mu^k - \mu^*||^2 - 2\alpha^k \left(q^* - q(\mu^k) + \sum_{i=1}^{m} \big(q_i(\mu^k) - q_i(\psi^{i-1,k})\big) \right) + m(\alpha^k)^2 C_i^2 \\ &\leq ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q^* - q(\mu^k)\big) + 2\alpha^k \sum_{i=2}^{m} C_i ||\psi^{i-1,k} - \mu^k|| + m(\alpha^k)^2 C_i^2 \\ &\leq ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q^* - q(\mu^k)\big) + (\alpha^k)^2 \left(2 \sum_{i=2}^{m} C_i \sum_{j=1}^{i-1} C_j + \sum_{i=1}^{m} C_i^2 \right) \\ &= ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q(\mu^*) - q(\mu^k)\big) + (\alpha^k)^2 \left(\sum_{i=1}^{m} C_i^2 \right)^2, \end{aligned}$$

where the next-to-last inequality follows from Eq. (2). Therefore

$$||\mu^{k+1} - \mu^*||^2 \leq ||\mu^k - \mu^*||^2 - 2\alpha^k \big(q^* - q(\mu^k)\big) + (\alpha^k)^2 C^2, \quad \forall\ \mu^* \in M^*, \quad \forall\ k \geq 0.$$

Assume that $\mu^k \notin M^*$ for all k. By substituting the expression for α^k in the above relation, we obtain

$$\begin{aligned} ||\mu^{k+1} - \mu^*||^2 &\leq ||\mu^k - \mu^*||^2 - \gamma^k (2 - \gamma^k) \frac{\big(q^* - q(\mu^k)\big)^2}{C^2} \\ &\leq ||\mu^k - \mu^*||^2 - \gamma_l (2 - \gamma_u) \frac{\big(q^* - q(\mu^k)\big)^2}{C^2}, \qquad \forall\ k \geq 0, \quad \forall\ \mu^* \in M^*. \end{aligned} \tag{3}$$

Therefore

$$||\mu^{k+1} - \mu^*|| < ||\mu^k - \mu^*||, \tag{4}$$

and the sequence $\{\mu^k\}$ is bounded. Next we will show that every limit point of $\{\mu^k\}$ belongs to M^*. Let $\{\mu^{k_j}\} \subset \{\mu^k\}$ and let $\bar{\mu}$ be such that $\lim_{j\to\infty} ||\mu^{k_j} - \bar{\mu}|| = 0$. Since the set M is closed, we have $\bar{\mu} \in M$. Suppose that $q(\bar{\mu}) < q^*$, i.e., $\bar{\mu} \notin M^*$. Since q is continuous, we can find a scalar $\delta > 0$ and an index j_0 such that

$$q(\mu^{k_j}) < q^* - \delta, \qquad \forall\ j \geq j_0.$$

This, combined with Eqs. (3) and (4), implies that

$$||\mu^{k_{j+1}} - \mu^*||^2 \leq ||\mu^{k_j} - \mu^*||^2 - \frac{\gamma_l (2 - \gamma_u) \delta^2}{C^2} \leq \cdots \leq ||\mu^{k_{j_0}} - \mu^*||^2 - (j + 1 - j_0)\ \frac{\gamma_l (2 - \gamma_u) \delta^2}{C^2},$$

which is a contradiction. Hence $\bar{\mu} \in M^*$. Note that the sequence of norms $\{||\mu^k - \mu^*||\}$ is strictly decreasing for any $\mu^* \in M^*$, so for any μ^* it converges to $||\bar{\mu} - \mu^*||$. Finally, to show that $\{\mu^k\}$ has a unique limit point, note that if $\hat{\mu} \in M^*$ and $\bar{\mu} \in M^*$ are limit points of the sequence $\{\mu^k\}$, we would have $||\bar{\mu} - \mu^*|| = ||\hat{\mu} - \mu^*||$ for all $\mu^* \in M^*$, which is possible only if $\hat{\mu} = \bar{\mu}$. This completes the proof.

6.3.17 (www)

For the separable problem

$$\begin{aligned}&\text{minimize} \quad \sum_{i=1}^{n} f_i(x_i)\\&\text{subject to} \quad \sum_{i=1}^{n} g_{ij}(x_i) \le 0, \quad j = 1, \ldots, r, \qquad \alpha_i \le x_i \le \beta_i, \quad i = 1, \ldots, n,\end{aligned}$$

where $f_i : \Re \mapsto \Re$, $g_{ij} : \Re \mapsto \Re$ are convex functions, the dual function is

$$q(\mu) = \sum_{i=1}^{n} \min_{\alpha_i \le x_i \le \beta_i} \left\{ f_i(x_i) + \sum_{j=1}^{r} \mu_j g_{ij}(x_i) \right\}.$$

Let $(\bar{x}, \bar{\mu})$ satisfy ϵ-complementary slackness as defined in the problem statement, and let $g(\bar{x})$ be the r-dimensional vector with jth component $\sum_{i=1}^{n} g_{ij}(\bar{x}_i)$. We will show that

$$q(\mu) \le q(\bar{\mu}) + \bar{\epsilon} + g(\bar{x})'(\mu - \bar{\mu}), \qquad \forall\ \mu \in \Re^r,$$

where

$$\bar{\epsilon} = \epsilon \sum_{i=1}^{n} (\beta_i - \alpha_i).$$

Indeed, we have for any $\mu \in \Re^r$

$$\begin{aligned} q(\mu) &\le \sum_{i=1}^{n} \left\{ f_i(\bar{x}_i) + \sum_{j=1}^{r} \mu_j g_{ij}(\bar{x}_i) \right\} \\ &= \sum_{i=1}^{n} \left\{ f_i(\bar{x}_i) + \sum_{j=1}^{r} \bar{\mu}_j g_{ij}(\bar{x}_i) \right\} + \sum_{j=1}^{r} (\mu_j - \bar{\mu}_j) \sum_{i=1}^{n} g_{ij}(\bar{x}_i) \qquad (1) \\ &= \sum_{i=1}^{n} \left\{ f_i(\bar{x}_i) + \sum_{j=1}^{r} \bar{\mu}_j g_{ij}(\bar{x}_i) \right\} + g(\bar{x})'(\mu - \bar{\mu}). \end{aligned}$$

For all i and all $x_i \in [\alpha_i, \beta_i]$, we have from the properties of directional derivatives and the convexity of the function $f_i(x_i) + \sum_{j=1}^{r} \bar{\mu}_j g_{ij}(x_i)$,

$$f_i(x_i) + \sum_{j=1}^{r} \bar{\mu}_j g_{ij}(x_i) \ge f_i(\bar{x}_i) + \sum_{j=1}^{r} \bar{\mu}_j g_{ij}(\bar{x}_i) + \gamma_i(x_i), \qquad (2)$$

where

$$\gamma_i(x_i) = \begin{cases} d_i^+(x_i - \bar{x}_i) & \text{if } \bar{x}_i = \alpha_i, \\ d_i^-(x_i - \bar{x}_i) & \text{if } \bar{x}_i = \beta_i, \\ \max\big\{ d_i^-(x_i - \bar{x}_i), d_i^+(x_i - \bar{x}_i) \big\} & \text{if } \alpha_i < \bar{x}_i < \beta_i, \end{cases} \qquad (3)$$

and

$$d_i^- = f_i^-(\bar{x}_i) + \sum_{j=1}^r \bar{\mu}_j g_{ij}^-(\bar{x}_i), \qquad d_i^+ = f_i^+(\bar{x}_i) + \sum_{j=1}^r \bar{\mu}_j g_{ij}^+(\bar{x}_i)$$

are the left and right derivatives of $f_i + \sum_{j=1}^r \bar{\mu}_j g_{ij}$ at $\bar{x}_i$. Using the ϵ-complementary slackness definition, we have

$$-\epsilon \le d_i^+ \quad \text{if} \quad \bar{x}_i = \alpha_i,$$

$$d_i^- \le \epsilon \quad \text{if} \quad \bar{x}_i = \beta_i,$$

$$-\epsilon \le d_i^- \le d_i^+ \le \epsilon \quad \text{if} \quad \alpha_i < \bar{x}_i < \beta_i.$$

Using the above relations in Eq. (3), we see that

$$\gamma_i(x_i) \ge -\epsilon(\beta_i - \alpha_i).$$

so Eq. (2) yields

$$f_i(x_i) + \sum_{j=1}^r \bar{\mu}_j g_{ij}(x_i) \ge f_i(\bar{x}_i) + \sum_{j=1}^r \bar{\mu}_j g_{ij}(\bar{x}_i) - \epsilon(\beta_i - \alpha_i), \quad \forall\ i,\ \forall\ x_i \in [\alpha_i, \beta_i].$$

By minimizing over $x_i \in [\alpha_i, \beta_i]$ and adding over i, and using the definition of the dual function $q(\bar{\mu})$, we obtain

$$\sum_{i=1}^n \left\{ f_i(\bar{x}_i) + \sum_{j=1}^r \bar{\mu}_j g_{ij}(\bar{x}_i) \right\} \le q(\bar{\mu}) + \epsilon \sum_{i=1}^n (\beta_i - \alpha_i),$$

which combined with Eq. (1), yields the desired relation

$$q(\mu) \le q(\bar{\mu}) + \bar{\epsilon} + g(\bar{x})'(\mu - \bar{\mu}).$$

6.3.18 www

For any $\mu \in M$, let us denote

$$d(\mu, M^*) = \min_{\mu \in M^*} \|\mu - \mu^*\|.$$

We first show that for all k, we have

$$\big(d(\mu^{k+1}, M^*)\big)^2 \le \big(d(\mu^k, M^*)\big)^2 - \frac{2s^k(\gamma - \beta)}{\gamma}\big(q^* - q(\mu^k)\big) + (s^k)^2(\delta + \beta)^2. \tag{1}$$

Indeed, using the definition of μ^{k+1}, the nonexpansive property of projection, the subgradient inequality, and the assumptions $\|r^k\| \le \beta$ and $\|g^k\| \le \delta$, we have for all $\mu^* \in M^*$,

$$\begin{aligned}
\big(d(\mu^{k+1}, M^*)\big)^2 &\le \|\mu^{k+1} - \mu^*\|^2 \\
&= \big\|\mu^k - \mu^* + s^k(g^k + r^k)\big\|^2 \\
&\le \|\mu^k - \mu^*\|^2 + 2s^k(g^k + r^k)'(\mu^k - \mu^*) + (s^k)^2\|g^k + r^k\|^2 \\
&\le \|\mu^k - \mu^*\|^2 + 2s^k g^{k\prime}(\mu^k - \mu^*) + 2s^k\|r^k\|\,\|\mu^k - \mu^*\| + (s^k)^2\|g^k + r^k\|^2 \\
&\le \|\mu^k - \mu^*\|^2 - 2s^k\big(q^* - q(\mu^k)\big) + 2s^k\beta\|\mu^k - \mu^*\| + (s^k)^2(\delta + \beta)^2.
\end{aligned}$$

If we let μ^* be the projection of μ^k on M^*, and use the assumption

$$q^* - q(\mu^k) \leq \gamma \min_{\mu^* \in M^*} \|\mu - \mu^*\|,$$

we obtain the desired relation (1).

Consider the case of a constant stepsize,

$$s^k = s, \qquad \forall\ k = 0, 1, \ldots$$

and to arrive at a contradiction, assume that for some nonnegative integer $\bar{k}$ and some $\epsilon > 0$, we have

$$\frac{s\gamma(\delta + \beta)^2}{2(\gamma - \beta)} + \epsilon < q^* - q(\mu^k), \qquad \forall\ k \geq \bar{k}. \tag{2}$$

Applying Eq. (1) with $s^k = s$, we obtain

$$\big(d(\mu^{k+1}, M^*)\big)^2 \leq \big(d(\mu^k, M^*)\big)^2 - \frac{2s(\gamma - \beta)}{\gamma}\big(q^* - q(\mu^k)\big) + s^2(\delta + \beta)^2,$$

which combined with Eq. (2), yields

$$\big(d(\mu^{k+1}, M^*)\big)^2 \leq \big(d(\mu^k, M^*)\big)^2 - \frac{2s(\gamma - \beta)}{\gamma}\left(\frac{s\gamma(\delta + \beta)^2}{2(\gamma - \beta)} + \epsilon\right) + s^2(\delta + \beta)^2$$

or

$$\big(d(\mu^{k+1}, M^*)\big)^2 \leq \big(d(\mu^k, M^*)\big)^2 - \frac{2s(\gamma - \beta)}{\gamma}\,\epsilon, \qquad \forall\ k \geq \bar{k}.$$

Since $\gamma > \beta$, this relation cannot hold for infinitely many k, thereby arriving at a contradiction.

The proof that $\limsup_{k \to \infty} q(\mu^k) = q^*$ is similar. To arrive at a contradiction, we assume that for some nonnegative integer $\bar{k}$ and some $\epsilon > 0$, we have

$$\epsilon < q^* - q(\mu^k), \qquad \forall\ k \geq \bar{k},$$

and we apply Eq. (1) to obtain

$$\big(d(\mu^{k+1}, M^*)\big)^2 \leq \big(d(\mu^k, M^*)\big)^2 - \frac{2s^k(\gamma - \beta)}{\gamma}\,\epsilon + (s^k)^2(\delta + \beta)^2, \qquad \forall\ k \geq \bar{k}.$$

Since $s^k \to 0$ and $\sum_{k=0}^{\infty} s^k = \infty$, this is a contradiction.