

Nonlinear Programming
3rd Edition

Theoretical Solutions Manual

Chapter 3


Dimitri P. Bertsekas

Massachusetts Institute of Technology



Athena Scientific, Belmont, Massachusetts

NOTE

This manual contains solutions of the theoretical problems, marked in the book by  It is continuously updated and improved, and it is posted on the internet at the book's www page

<http://www.athenasc.com/nonlinbook.html>

Many thanks are due to several people who have contributed solutions, and particularly to David Brown, Angelia Nedic, Asuman Ozdaglar, Cynara Wu.

Last Updated: May 2016

Solutions Chapter 3

SECTION 3.1

3.1.9 www

From Prop. 3.1.2(a), if x^* is a local minimum, then

$$\nabla f(x^*)'(x - x^*) \geq 0, \quad \forall x \in X,$$

or

$$\sum_{i=1}^n \frac{\partial f(x^*)}{\partial x_i} (x_i - x_i^*) \geq 0.$$

If $x_i^* = \alpha_i$, then $x_i \geq x_i^*, \forall x_i$. Letting $x_j = x_j^*$, for $j \neq i$, we have

$$\frac{\partial f(x^*)}{\partial x_i} \geq 0.$$

Similarly, if $x_i^* = \beta_i$, then $x_i \leq x_i^*$, for all x_i . Letting $x_j = x_j^*$, for $j \neq i$, we have

$$\frac{\partial f(x^*)}{\partial x_i} \leq 0.$$

If $\alpha_i < x_i^* < \beta_i$, let $x_j = x_j^*$ for $j \neq i$. Letting $x_i = \alpha_i$, we obtain

$$\frac{\partial f(x^*)}{\partial x_i} \leq 0,$$

and letting $x_i = \beta_i$, we obtain

$$\frac{\partial f(x^*)}{\partial x_i} \geq 0.$$

Combining these inequalities, we see that we must have

$$\frac{\partial f(x^*)}{\partial x_i} = 0.$$

Assume that f is convex. To show that Eqs. (1.6)-(1.8) are sufficient for x^* to be a global minimum, let $I_1 = \{i \mid x_i^* = \alpha_i\}$, $I_2 = \{i \mid x_i^* = \beta_i\}$, $I_3 = \{i \mid \alpha_i < x_i^* < \beta_i\}$. Then

$$\begin{aligned} \nabla f(x^*)'(x - x^*) &= \sum_{i=1}^n \frac{\partial f(x^*)}{\partial x_i} (x_i - x_i^*) \\ &= \sum_{i \in I_1} \frac{\partial f(x^*)}{\partial x_i} (x_i - \alpha_i) + \sum_{i \in I_2} \frac{\partial f(x^*)}{\partial x_i} (x_i - \beta_i) + \sum_{i \in I_3} \frac{\partial f(x^*)}{\partial x_i} (x_i - x_i^*). \end{aligned}$$

Since $\frac{\partial f(x^*)}{\partial x_i} \geq 0$ for $i \in I_1$, $\frac{\partial f(x^*)}{\partial x_i} \leq 0$ for $i \in I_2$, and $\frac{\partial f(x^*)}{\partial x_i} = 0$ for $i \in I_3$, each term in the above equation is greater than or equal to zero. Therefore

$$\nabla f(x^*)'(x - x^*) \geq 0, \quad \forall x \in X.$$

From Prop. 3.1.2(b), it follows that x^* is a global minimum.

3.1.10 www

For any $x \in X$ such that $\nabla f(x^*)'(x - x^*) = 0$, we have by the second order expansion of Prop. A.23, for all $\alpha \in [0, 1]$ and some $\tilde{\alpha} \in [0, \alpha]$,

$$f(x^* + \alpha(x - x^*)) - f(x^*) = \frac{1}{2}\alpha^2(x - x^*)'\nabla^2 f(x^* + \tilde{\alpha}(x - x^*))(x - x^*).$$

For all sufficiently small α , the left-hand side is nonnegative, since x^* is a local minimum. Hence the same is true for the right-hand side, and by taking the limit as $\alpha \rightarrow 0$ (and also $\tilde{\alpha} \rightarrow 0$), we obtain

$$(x - x^*)'\nabla^2 f(x^*)(x - x^*) \geq 0.$$

3.1.11 www

Proof under condition (1): Assume, to arrive at a contradiction, that x^* is not a local minimum. Then there exists a sequence $\{x^k\} \subseteq X$ converging to x^* such that $f(x^k) < f(x^*)$ for all k . We have

$$f(x^k) = f(x^*) + \nabla f(x^*)'(x^k - x^*) + \frac{1}{2}(x^k - x^*)'\nabla^2 f(x^*)(x^k - x^*) + o(\|x^k - x^*\|^2).$$

Introducing the vector

$$p^k = \frac{x^k - x^*}{\|x^k - x^*\|},$$

and using the relation $f(x^k) < f(x^*)$, we obtain

$$\nabla f(x^*)'p^k + \frac{1}{2}p^{k'}\nabla^2 f(x^*)p^k\|x^k - x^*\| + \frac{o(\|x^k - x^*\|^2)}{\|x^k - x^*\|} < 0. \quad (1)$$

This together with the hypothesis $\nabla f(x^*)'p^k \geq 0$ implies

$$\frac{1}{2}p^{k'}\nabla^2 f(x^*)p^k\|x^k - x^*\| + \frac{o(\|x^k - x^*\|^2)}{\|x^k - x^*\|} < 0. \quad (2)$$

Let us call *feasible direction at x^** any vector p of the form $\alpha(x - x^*)$, where $\alpha > 0$ and $x \in X$, $x \neq x^*$ (see also Section 3.2). The sequence $\{p^k\}$ is a sequence of feasible directions at

x^* that lie on the surface of the unit sphere. Therefore, a subsequence $\{p^k\}_K$ converges to a vector \bar{p} , which because X is polyhedral, must be a feasible direction at x^* (this is easily seen by expressing the polyhedral set X in terms of linear equalities and inequalities). Therefore, by the hypothesis of the exercise, we have $\nabla f(x^*)'\bar{p} \geq 0$. By letting $k \rightarrow \infty$, $k \in K$ in (1), we have

$$\nabla f(x^*)'\bar{p} = 0.$$

The hypothesis of the exercise implies that

$$\bar{p}'\nabla^2 f(x^*)\bar{p} > 0. \quad (3)$$

Dividing by $\|x^k - x^*\|$ and taking the limit in Eq. (2) as $k \rightarrow \infty$, $k \in K$, we obtain

$$\frac{1}{2}\bar{p}'\nabla^2 f(x^*)\bar{p} + \lim_{k \rightarrow \infty, k \in K} \frac{o(\|x^k - x^*\|^2)}{\|x^k - x^*\|^2} \leq 0.$$

This contradicts Eq. (3).

Proof under condition (2): Here we argue in the similar way as in part (1). Suppose that all the given assumptions hold and x^* is not a local minimum. Then there is a sequence $\{x^k\} \subseteq X$ converging to x^* such that $f(x^k) < f(x^*)$ for all k . By using the second order expansion of f at x^* and introducing the vector $p^k = \frac{x^k - x^*}{\|x^k - x^*\|}$, we have that both Eq. (1) and (2) hold for all k . Since $\{p^k\}$ consists of feasible directions at x^* that lie on the surface of the unit sphere, there is a subsequence $\{p^k\}_K$ converging to a vector \bar{p} with $\|\bar{p}\| = 1$. By the assumption given in the exercise, we have that

$$\nabla f(x^*)p^k \geq 0, \quad \forall k.$$

Hence $\nabla f(x^*)\bar{p} \geq 0$. By letting $k \rightarrow \infty$, $k \in K$ in (1), we obtain $\nabla f(x^*)\bar{p} \leq 0$. Consequently $\nabla f(x^*)\bar{p} = 0$. Since the vector \bar{p} is in the closure of the set of the feasible directions at x^* , the condition given in part (2) implies that $\bar{p}'\nabla^2 f(x^*)\bar{p} > 0$. Dividing by $\|x^k - x^*\|$ and taking the limit in Eq. (2) as $k \rightarrow \infty$, $k \in K$, we obtain $\bar{p}'\nabla^2 f(x^*)\bar{p} \leq 0$, which is a contradiction. Therefore, x^* must be a local minimum.

Proof under condition (3): We have

$$f(x) = f(x^*) + \nabla f(x^*)'(x - x^*) + \frac{1}{2}(x - x^*)'\nabla^2 f(x^*)(x - x^*) + o(\|x - x^*\|^2),$$

so that by using the hypotheses $\nabla f(x^*)'(x - x^*) \geq 0$ and $(x - x^*)'\nabla^2 f(x^*)(x - x^*) \geq \gamma\|x - x^*\|^2$,

$$f(x) - f(x^*) \geq \frac{\gamma}{2}\|x - x^*\|^2 + o(\|x - x^*\|^2).$$

The expression in the right-hand side is nonnegative for $x \in X$ close enough to x^* , and it is strictly positive if in addition $x \neq x^*$. Hence x^* is a strict local minimum.

Example: [Why the assumption that X is a polyhedral set was important under condition (1)] A polyhedral set X has the property that for any point $x \in X$, the set $V(x)$ of the feasible directions at x is closed. This was crucial for proving that the conditions

$$\nabla f(x^*)'(x - x^*) \geq 0, \quad \forall x \in X, \quad (1)$$

$$(x - x^*)\nabla^2 f(x^*)(x - x^*) > 0, \quad \forall x \in X, x \neq x^*, \text{ for which } \nabla f(x^*)'(x - x^*) = 0, \quad (2)$$

are sufficient for local optimality of x^* .

Consider the set $X = \{(x_1, x_2) \mid (x_1)^2 \leq x_2\}$ and the point $(0, 0) \in X$. Let the cost function be $f(x_1, x_2) = -2(x_1)^2 + x_2$. Note that the gradient of f at 0 is $[0, 1]'$. It is easy to see that

$$\nabla f(0)'(x - 0) = x_2 > 0, \quad \forall x \in X, x \neq 0.$$

Thus the point $x^* = 0$ satisfies conditions (1) and (2) (condition 2 is trivially satisfied since in our example $\nabla f(0)'(x - 0) = 0$ simply never occurs for $x \in X, x \neq 0$). On the other hand, $x^* = 0$ is not a local minimum of f in X . Consider the points $x^n = (\frac{1}{n}, \frac{1}{n^2}) \in X$ for $n \geq 1$. Since $x^n \rightarrow x^*$ as $n \rightarrow \infty$, for any $\delta > 0$ there is an index n_δ such that $\|x^n - x^*\| < \delta$ for all $n \geq n_\delta$. By evaluating the cost function, we have $f(x^n) = -\frac{1}{n^2} < 0 = f(x^*)$. Hence, in any δ neighborhood of $x^* = 0$, there are points $x^n \in X$ with the better objective value, i.e. x^* is not a local minimum.

This is happening because the set $V(x^*)$ of the feasible directions at point x^* is not closed in this case. The set $V(x^*)$ is given by

$$V(x^*) = \{d = (d_1, d_2) \mid d_2 > 0, \|d\| = 1\},$$

and is open. The vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

belong to the closure of $V(x^*)$ but they are not in the set $V(x^*)$.

3.1.18 www

The assumption on $\nabla^2 f(x)$ guarantees that f is strictly convex and coercive, so it has a unique global minimum over any closed convex set (using Weierstrass' theorem, Prop. A.8). By the second order expansion of Prop. A.23, we have for all x and y in \mathfrak{R}^n

$$f(y) = f(x) + \nabla f(x)'(y - x) + \frac{1}{2}(y - x)\nabla^2 f(\tilde{y})(y - x)$$

for some \tilde{y} in the line segment connecting x and y . It follows, using the hypothesis, that

$$\nabla f(x)'(y-x) + \frac{M}{2}\|y-x\|^2 \geq f(y) - f(x) \geq \nabla f(x)'(y-x) + \frac{m}{2}\|y-x\|^2.$$

Taking the minimum in this inequality over $y \in X$, and changing sign, we obtain

$$-\min_{y \in X} \left\{ \nabla f(x)'(y-x) + \frac{M}{2}\|y-x\|^2 \right\} \leq f(x) - f(x^*) \leq -\min_{y \in X} \left\{ \nabla f(x)'(y-x) + \frac{m}{2}\|y-x\|^2 \right\},$$

which is the desired relation.

3.1.19 (Existence of Solutions of Nonconvex Quadratic Programming

Problems) www

Let $\{\gamma^k\}$ be a decreasing sequence with $\gamma^k \downarrow f^*$, and denote

$$S^k = \{x \in X \mid x'Qx + c'x \leq \gamma^k\}.$$

Then the set of optimal solutions of the problem is $\bigcap_{k=0}^{\infty} S^k$, so by Prop. 3.1.4, it will suffice to show that for each asymptotic direction of $\{S^k\}$, all corresponding asymptotic sequences are retractive. Let d be an asymptotic direction and let $\{x^k\}$ be a corresponding asymptotic sequence. Similar to the proof of Prop. 3.1.5, we have $d'Qd \leq 0$. Also, in case (i), similar to the proof of Prop. 3.1.5, we have $a'_j d \leq 0$ for all j , while in case (ii) it is seen that $d \in N$, where $X = B + N$ and B is compact and N is a polyhedral cone. For any $x \in X$, consider the vectors $\tilde{x}^k = x + kd$. Then, in both cases (i) and (ii), it can be seen that $\tilde{x}^k \in X$ [in case (i) by using the argument in the proof of Prop. 3.1.5, and in case (ii) by using the definition $X = B + N$]. Thus, the cost function value corresponding to \tilde{x}^k satisfies

$$\begin{aligned} f^* &\leq (x + kd)'Q(x + kd) + c'(x + kd) \\ &= x'Qx + c'x + k^2 d'Qd + k(c + 2Qx)'d \\ &\leq x'Qx + c'x + k(c + 2Qx)'d, \end{aligned}$$

where the last inequality follows from the fact $d'Qd \leq 0$. From the finiteness of f^* , it follows that

$$(c + 2Qx)'d \geq 0, \quad \forall x \in X.$$

We now show that $\{x^k\}$ is retractive, so that we can use Prop. 3.1.4. Indeed for any $\alpha > 0$, since $\|x^k\| \rightarrow \infty$, it follows that for k sufficiently large, we have $x^k - \alpha d \in X$ [this follows similar to the proof of Prop. 3.1.5 in case (i), and because $d \in N$ in case (ii)]. Furthermore, we have

$$\begin{aligned} f(x^k - \alpha d) &= (x^k - \alpha d)'Q(x^k - \alpha d) + c'(x^k - \alpha d) \\ &= x^{k'}Qx^k + c'x^k - \alpha(c + 2Qx^k)'d + \alpha^2 d'Qd \\ &\leq x^{k'}Qx^k + c'x^k \\ &\leq \gamma^k, \end{aligned}$$

where the first inequality follows from the facts $d'Qd \leq 0$ and $(c + 2Qx^k)'d \geq 0$ shown earlier. Thus for sufficiently large k , we have $x^k - \alpha d \in S^k$, so that $\{x^k\}$ is retractive. The existence of an optimal solution now follows from Prop. 3.1.4.

3.1.20 www

We proceed as in the proof of Prop. 3.1.5. By using a decomposition of d^k as the sum of a vector in the nullspace of A and its orthogonal complement, and an argument like the one in the proof of Prop. 3.1.5, we can show that

$$Ad = 0, \quad c'd \leq 0.$$

Similarly, we can show that

$$a'_j d \leq 0, \quad j = 1, \dots, r.$$

Using the finiteness of f^* , we can also show that $c'd = 0$, and we can conclude the proof similar to the proof of Prop. 3.1.5.

3.1.21 www

Note that the cone N in this exercise must be assumed polyhedral (see the errata sheet). Let $S^k = \{x \in X \mid f(x) \leq \gamma^k\}$, and let d be an asymptotic direction of $\{S^k\}$, and let $\{x^k\}$ be a corresponding asymptotic sequence. We will show that $\{x^k\}$ is retractive, so by applying Prop. 3.1.4, it follows that the intersection of $\{S^k\}$, the set of minima of f over X , is nonempty.

Since d is an asymptotic direction of $\{S^k\}$, d is also an asymptotic direction of $\{x \mid f(x) \leq \gamma^k\}$, and by hypothesis for some bounded positive sequence $\{\alpha^k\}$ and some positive integer \bar{k} , we have $f(x^k - \alpha^k d) \leq \gamma^k$ for all $k \geq \bar{k}$.

Let $X = \bar{X} + N$, where \bar{X} is compact, and N is the polyhedral cone

$$N = \{y \mid a'_j y \leq 0, j = 1, \dots, r\},$$

where a_1, \dots, a_r are some vectors. We can represent x^k as

$$x^k = \bar{x}^k + y^k, \quad \forall k = 0, 1, \dots,$$

where $\bar{x}^k \in \bar{X}$ and $y^k \in N$, so that

$$a'_j x^k = a'_j (\bar{x}^k + y^k), \quad \forall k = 0, 1, \dots, j = 1, \dots, r.$$

Dividing both sides with $\|x^k\|$ and taking the limit as $k \rightarrow \infty$, we obtain

$$a'_j d = \lim_{k \rightarrow \infty} \frac{a'_j y^k}{\|x^k\|}.$$

Since $a'_j y^k \leq 0$ for all k and j , we obtain that $a'_j d \leq 0$ for all j , so that $d \in N$.

For each j , we consider two cases:

- (1) $a'_j d = 0$. In this case, $a'_j(y^k - \bar{\alpha}d) \leq 0$ for all k , since $y^k \in N$ and $a'_j y^k \leq 0$.
- (2) $a'_j d < 0$. In this case, we have

$$\frac{1}{\|x^k\|} a'_j(y^k - \bar{\alpha}d) = \frac{1}{\|x^k\|} a'_j(x^k - \bar{x}^k - \bar{\alpha}d),$$

so that since $\frac{x^k}{\|x^k\|} \rightarrow d$, $\{x^k\}$ is unbounded, and $\{\bar{x}^k\}$ is bounded, we obtain

$$\lim_{k \rightarrow \infty} \frac{1}{\|x^k\|} a'_j(y^k - \bar{\alpha}d) = a'_j d < 0.$$

Hence $a'_j(y^k - \bar{\alpha}d) < 0$ for k greater than some \bar{k} .

Thus, for $k \geq \bar{k}$ and $\alpha \in (0, \bar{\alpha}]$, we have $a'_j(y^k - \alpha d) \leq a'_j(y^k - \bar{\alpha}d) \leq 0$ for all j , so that $y^k - \alpha d \in N$ and $x^k - \alpha d \in X$.

Thus $\{x^k\}$ is retractive, and by applying Prop. 3.1.4, we have that $\{S^k\}$ has nonempty intersection.

3.1.22 www

We follow the hint. Let $\{y_k\}$ be a sequence of points in AS converging to some $\bar{y} \in \mathfrak{R}^n$. We will prove that AS is closed by showing that $\bar{y} \in AS$.

We introduce the sets

$$W_k = \{z \mid \|z - \bar{y}\| \leq \|y_k - \bar{y}\|\},$$

and

$$S_k = \{x \in S \mid Ax \in W_k\}.$$

To show that $\bar{y} \in AS$, it is sufficient to prove that the intersection $\bigcap_{k=0}^{\infty} S_k$ is nonempty, since every $\bar{x} \in \bigcap_{k=0}^{\infty} S_k$ satisfies $\bar{x} \in S$ and $A\bar{x} = \bar{y}$ (because $y_k \rightarrow \bar{y}$). The asymptotic directions of $\{S_k\}$ are asymptotic directions of S that are also in the nullspace of A , and it can be seen that every corresponding asymptotic sequence is retractive for $\{S_k\}$. Hence, by Prop. 3.1.4, $\bigcap_{k=0}^{\infty} S_k$ is nonempty.

SECTION 3.2

3.2.7 www

Since the number of extreme points of f is finite, some extreme point must be repeated within a finite number of iterations, i.e., for some k and $i \in \{0, 1, \dots, k-1\}$, we have

$$\bar{x}^i = \arg \min_{x \in X} \nabla f(x^k)'(x - x^k).$$

Since x^k minimizes $f(x)$ over X^{k-1} , we must have

$$\nabla f(x^k)'(\bar{x}^i - x^k) \geq 0, \quad \forall i = 0, 1, \dots, k-1.$$

Combining the above two equations, we see that

$$\nabla f(x^k)'(x - x^k) \geq 0, \quad \forall x \in X,$$

which implies that x^k is a stationary point of f over X .

SECTION 3.3

3.3.4 www

We assume here that the unscaled version of the method ($H^k = I$) is used and that the stepsize s^k is a constant $s > 0$.

(a) If x^k is nonstationary, there exists a feasible descent direction $\hat{x}^k - x^k$ for the original problem, where $\hat{x}^k \in X$. Since $\hat{x}^k \in X^k$, we have

$$\nabla f(x^k)'(\tilde{x}^k - x^k) + \frac{1}{2s} \|\tilde{x}^k - x^k\|^2 \leq \nabla f(x^k)'(\hat{x}^k - x^k) + \frac{1}{2s} \|\hat{x}^k - x^k\|^2 < 0,$$

where \tilde{x}^k is defined by the algorithm. Thus,

$$\nabla f(x^k)'(\tilde{x}^k - x^k) \leq -\frac{1}{2s} \|\tilde{x}^k - x^k\|^2 < 0,$$

so that $\tilde{x}^k - x^k$ is a descent direction at x^k . It is also a feasible direction, since $a_j'(\tilde{x}^k - x^k) \leq 0$ for all j such that $a_j x^k = b_j$.

(b) As in the proof of Prop. 3.3.1, we will show that the direction sequence $\{\bar{x}^k - x^k\}$ is gradient-related, where

$$\bar{x}^k = \gamma^k \tilde{x}^k + (1 - \gamma^k) x^k$$

and

$$\gamma^k = \max \left\{ \gamma \in [0, 1] \mid \gamma \tilde{x}^k + (1 - \gamma)x^k \in X \right\}.$$

Indeed, suppose that $\{x^k\}_{k \in K}$ converges to a nonstationary point \tilde{x} . We must prove that

$$\limsup_{k \rightarrow \infty, k \in K} \|\bar{x}^k - x^k\| < \infty, \quad (*)$$

$$\limsup_{k \rightarrow \infty, k \in K} \nabla f(x^k)'(\bar{x}^k - x^k) < 0. \quad (**)$$

Since $\|\bar{x}^k - x^k\| \leq \|\tilde{x}^k - x^k\| \leq s \|\nabla f(x^k)\|$, Eq. (*) clearly holds, so we concentrate on proving (**). The key to this is showing that γ^k is bounded away from 0, so that the inner product $\nabla f(x^k)'(\bar{x}^k - x^k)$ is bounded away from 0 when $\nabla f(x^k)'(\tilde{x}^k - x^k)$ is.

For each k , we either have $\gamma^k = 1$, or else we must have for some j with $a'_j x^k < b_j - \epsilon$,

$$a'_j (\gamma^k \tilde{x}^k + (1 - \gamma^k)x^k) = b_j$$

so that

$$\gamma^k a'_j (\tilde{x}^k - x^k) = b_j - a'_j x^k > \epsilon,$$

from which

$$\gamma^k > \frac{\epsilon}{\|a_j\| \cdot \|\tilde{x}^k - x^k\|}.$$

It follows that for all k , we have

$$\min \left\{ 1, \min_j \frac{\epsilon}{\|a_j\| \cdot \|\tilde{x}^k - x^k\|} \right\} \leq \gamma^k \leq 1.$$

Since the subsequence $\{x^k\}_K$ converges, the subsequence $\{\tilde{x}^k - x^k\}_K$ is bounded implying also that the subsequence $\{\gamma^k\}_K$ is bounded away from 0.

For sufficiently large k , the set

$$X^k = \{x \mid a'_j x \leq b_j, \text{ for all } j \text{ with } b_j - \epsilon \leq a'_j x^k \leq b_j\},$$

is equal to the set

$$\tilde{X} = \{x \mid a'_j x \leq b_j, \text{ for all } j \text{ with } b_j - \epsilon \leq a'_j \tilde{x} \leq b_j\},$$

so proceeding as in the proof of Prop. 3.3.1, we obtain

$$\limsup_{k \rightarrow \infty, k \in K} \nabla f(x^k)'(\tilde{x}^k - x^k) \leq -\frac{1}{s} \|\tilde{x} - [\tilde{x} - s \nabla f(\tilde{x})]^+\|^2,$$

where $[\cdot]^+$ denotes projection on the set \tilde{X} . Since \tilde{x} is nonstationary, the right-hand side of the above inequality is negative, so that

$$\limsup_{k \rightarrow \infty, k \in K} \nabla f(x^k)'(\tilde{x}^k - x^k) < 0.$$

We have $\bar{x}^k - x^k = \gamma^k(\tilde{x}^k - x^k)$, and since γ^k is bounded away from 0, it follows that

$$\limsup_{k \rightarrow \infty, k \in K} \nabla f(x^k)'(\bar{x}^k - x^k) < 0,$$

proving Eq. (**).

(c) Here we consider the variant of the method that uses a constant stepsize, which however, is reduced if necessary to ensure that \bar{x}^k is feasible. If the stepsize is sufficiently small to ensure convergence to the unique local minimum x^* of the positive definite quadratic cost function, then \bar{x}^k will be arbitrarily close to x^* for sufficiently large k , so that $\bar{x}^k = \tilde{x}^k$. Thus the convergence rate estimate of the text applies.

3.3.7 www

The key idea is to show that x^k stays in the bounded set

$$A = \{x \in X \mid f(x) \leq f(x^0)\}$$

and to use a constant stepsize $s^k = s$ that depends on the constant L corresponding to this bounded set. Let

$$R = \max\{\|x\| \mid x \in A\},$$

$$G = \max\{\|\nabla f(x)\| \mid x \in A\},$$

and

$$B = \{x \mid \|x\| \leq R + 2G\}.$$

Using condition (i) in the exercise, there exists some constant L such that $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$, for all $x, y \in B$. Suppose the stepsize s satisfies $0 < s < 2 \min\{1, 1/L\}$. We will, show by induction on k that with this stepsize, we have $x^k \in A$ and

$$f(x^{k+1}) \leq f(x^k) - \left(\frac{L}{2} - \frac{1}{s}\right) \|x^{k+1} - x^k\|^2 \leq f(x^k), \quad (*)$$

for all $k \geq 0$.

To start the induction, we note that $x^0 \in A$, by the definition of A . Suppose that $x^k \in A$. We have $x^{k+1} = [x^k - s\nabla f(x^k)]^+$, so by using the nonexpansiveness of the projection mapping,

$$\|x^{k+1} - x^k\| \leq \|(x^k - s\nabla f(x^k)) - x^k\| \leq s\|\nabla f(x^k)\| \leq 2G.$$

Thus,

$$\|x^{k+1}\| \leq \|x^k\| + 2G \leq R + 2G,$$

implying that $x^{k+1} \in B$. Since B is convex, we conclude that the entire line segment connecting x^k and x^{k+1} belongs to B . In order to prove Eq. (*), we now proceed as in the proof of Prop. 3.3.2. A difficulty arises because Prop. A.24 assumes that the inequality $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ holds for all x, y , whereas in this exercise this inequality holds only for $x, y \in B$. However, using the fact that the Lipschitz condition holds along the line segment connecting x^k and x^{k+1} (which belongs to B as argued earlier), the proof of Prop. A.24 can be repeated to obtain

$$f(x^{k+1}) - f(x^k) \leq \nabla f(x^k)'(x^{k+1} - x^k) + \frac{L}{2}\|x^{k+1} - x^k\|^2.$$

Using this relation, and the relation

$$\nabla f(x^k)'(x^{k+1} - x^k) \leq -\frac{1}{s}\|x^{k+1} - x^k\|^2,$$

we obtain Eq. (*). It follows that $x^{k+1} \in A$, completing the induction. The remainder of the proof is the same as in Prop. 3.3.2.