Corrections for the book NONLINEAR PROGRAMMING: 2ND EDITION, Athena Scientific, 1999, by Dimitri P. Bertsekas

Note: The first set of corrections given below are for the 1st Printing of the book. Many of these corrections have been incorporated in the 2nd and 3rd Printings of the book. See the end of this file for additional corrections to the 2nd and 3rd Printings.

Last updated: 1/18/2016

Corrections to the 1ST PRINTING

p. 16 (-6) Change $x < 2\pi$ to $y < 2\pi$

p. 21 (+4) Change $\frac{4\pi}{3}$ to $\frac{5\pi}{6}$

p. 43 The following is a more streamlined proof of Prop. 1.2.1 (it eliminates the vector p^k). The modifications begin at the 8th line of p. 44, where p^k is introduced, but the proof is given here in its entirety for completeness.

Proof: Consider the Armijo rule, and to arrive at a contradiction, assume that $\overline{x}$ is a limit point of $\{x^k\}$ with $\nabla f(\overline{x}) \neq 0$. Note that since $\{f(x^k)\}$ is monotonically nonincreasing, $\{f(x^k)\}$ either converges to a finite value or diverges to $-\infty$. Since f is continuous, $f(\overline{x})$ is a limit point of $\{f(x^k)\}$, so it follows that the entire sequence $\{f(x^k)\}$ converges to $f(\overline{x})$. Hence,

$$f(x^k) - f(x^{k+1}) \to 0.$$

By the definition of the Armijo rule, we have

$$f(x^k) - f(x^{k+1}) \geq -\sigma\alpha^k\nabla f(x^k)'d^k. \tag{1.16}$$

Hence, $\alpha^k\nabla f(x^k)'d^k \to 0$. Let $\{x^k\}_{\mathcal{K}}$ be a subsequence converging to $\overline{x}$. Since $\{d^k\}$ is gradient related, we have

$$\limsup_{\substack{k\to\infty \\ k\in\mathcal{K}}} \nabla f(x^k)'d^k < 0,$$

and therefore

$$\{\alpha^k\}_{\mathcal{K}} \to 0.$$

Hence, by the definition of the Armijo rule, we must have for some index $\overline{k} \geq 0$

$$f(x^k) - f\big(x^k + (\alpha^k/\beta)d^k\big) < -\sigma(\alpha^k/\beta)\nabla f(x^k)'d^k, \qquad \forall\ k \in \mathcal{K},\ k \geq \overline{k}, \tag{1.17}$$

that is, the initial stepsize s will be reduced at least once for all $k \in \mathcal{K}$, $k \geq \overline{k}$. Since $\{d^k\}$ is gradient related, $\{d^k\}_{\mathcal{K}}$ is bounded, and it follows that there exists a subsequence $\{d^k\}_{\overline{\mathcal{K}}}$ of $\{d^k\}_{\mathcal{K}}$ such that

$$\{d^k\}_{\overline{\mathcal{K}}} \to \overline{d},$$

where $\overline{d}$ is some vector which must be nonzero in view of the definition of a gradient related sequence. From Eq. (1.17), we have

$$\frac{f(x^k) - f(x^k + \overline{\alpha}^k d^k)}{\overline{\alpha}^k} < -\sigma \nabla f(x^k)' d^k, \qquad \forall\, k \in \overline{\mathcal{K}},\ k \geq \overline{k}, \tag{1.18}$$

where $\overline{\alpha}^k = \alpha^k/\beta$. By using the mean value theorem, this relation is written as

$$-\nabla f(x^k + \tilde{\alpha}^k d^k)' d^k < -\sigma \nabla f(x^k)' d^k, \qquad \forall\, k \in \overline{\mathcal{K}},\ k \geq \overline{k},$$

where $\tilde{\alpha}^k$ is a scalar in the interval $[0, \overline{\alpha}^k]$. Taking limits in the above equation we obtain

$$-\nabla f(\overline{x})' \overline{d} \leq -\sigma \nabla f(\overline{x})' \overline{d}$$

or

$$0 \leq (1 - \sigma) \nabla f(\overline{x})' \overline{d}.$$

Since $\sigma < 1$, it follows that

$$0 \leq \nabla f(\overline{x})' \overline{d}, \tag{1.19}$$

which contradicts the definition of a gradient related sequence. This proves the result for the Armijo rule.

Consider now the minimization rule, and let $\{x^k\}_{\mathcal{K}}$ converge to $\overline{x}$ with $\nabla f(\overline{x}) \neq 0$. Again we have that $\{f(x^k)\}$ decreases monotonically to $f(\overline{x})$. Let $\tilde{x}^{k+1}$ be the point generated from x^k via the Armijo rule, and let $\tilde{\alpha}^k$ be the corresponding stepsize. We have

$$f(x^k) - f(x^{k+1}) \geq f(x^k) - f(\tilde{x}^{k+1}) \geq -\sigma \tilde{\alpha}^k \nabla f(x^k)' d^k.$$

By repeating the arguments of the earlier proof following Eq. (1.16), replacing α^k by $\tilde{\alpha}^k$, we can obtain a contradiction. In particular, we have

$$\{\tilde{\alpha}^k\}_{\mathcal{K}} \to 0,$$

and by the definition of the Armijo rule, we have for some index $\overline{k} \geq 0$

$$f(x^k) - f\big(x^k + (\tilde{\alpha}^k/\beta) d^k\big) < -\sigma (\tilde{\alpha}^k/\beta) \nabla f(x^k)' d^k, \qquad \forall\, k \in \mathcal{K},\ k \geq \overline{k},$$

[cf. Eq. (1.17)]. Proceeding as earlier, we obtain Eqs. (1.18) and (1.19) (with $\overline{\alpha}^k = \tilde{\alpha}^k/\beta$), and a contradiction of Eq. (1.19).

The line of argument just used establishes that any stepsize rule that gives a larger reduction in cost at each iteration than the Armijo rule inherits its convergence properties. This also proves the proposition for the limited minimization rule. **Q.E.D.**

p. 49 (-4 and -3) Change $f(x^k)$ to $\nabla f(x^k)$

p. 54 (+11) Change "Condition (i)" to "The hypothesis"

p. 54 (-9) Change "condition (i)" to "the hypothesis"

p. 59 (Figure) Change "Eqs. (2.26) and (2.27)" to "Eqs. (1.34) and (1.35)"

p. 101 (+1) Change the first two equations to

$$c_1 \min\big\{\nabla f(x^k)' D \nabla f(x^k),\ \|\nabla f(x^k)\|^3\big\} \leq -\nabla f(x^k)'\big((1-\beta^m)d_S^k + \beta^m d_N\big),$$

$$\|(1-\beta^m)d_S^k + \beta^m d_N\| \leq c_2 \max\big\{\|D\nabla f(x^k)\|,\ \|\nabla f(x^k)\|^{1/2}\big\},$$

p. 116 (-5) Change "$\|x^k - x^*\|$" to "$\|x^k - x(\alpha)\|$"

p. 134 (-11) Change "be a linear combination" to "either be 0 or be a linear combination"

p. 156 (+14) Change $\frac{p^k p^{k\prime}}{p^{k\prime} p^k}$ to $\frac{p^k p^{k\prime}}{p^{k\prime} q^k}$

p. 162 (+20) Change "that is" to "i.e.,"

p. 176 (+2) Change "Prob. 1.8.1" to "Prob. 1.9.1"

p. 183 (-18) Change "is very important" to "are very important"

p. 188 (+23) Change [WiH59] to [WiH60]

p. 212 (-4) Change $\alpha^k = \beta^{m_k} s$ to $\alpha^k = \beta^{m_k}$

p. 221 (+5) Change x^0 to x^k

p. 221 (-6) Change x_0 to x^k

p. 227 (+15) Change "Zoutendijk's method uses ..." to "Zoutendijk's method rescales d_k so that $x_k + d_k$ is feasible, and uses ..."

p. 268 (+18) The assumption of Proposition 2.7.1 should be modified to include a condition that $f(x_1, \ldots, x_{i-1}, \xi, x_{i+1}, \ldots, x_m)$ viewed as a function of ξ, attains a unique minimum $\bar{\xi}$ over X_i, AND is monotonically nonincreasing in the interval from x_i to $\bar{\xi}$. [In the 1st edition of the book, the function $f(x_1, \ldots, x_{i-1}, \xi, x_{i+1}, \ldots, x_m)$ was assumed strictly convex in ξ, but in the process of generalizing the proposition in the 2nd edition, the assumptions were not stated correctly.] What follows is a corrected statement with an edited proof and a comment on the assumptions at the end. Replace Prop. 7.2.1 and its proof with the following:

Proposition 2.7.1: (Convergence of Block Coordinate Descent) Suppose that f is continuously differentiable over the set X of Eq. (2.111). Furthermore, suppose that for each $x = (x_1, \ldots, x_m) \in X$ and i,

$$f(x_1, \ldots, x_{i-1}, \xi, x_{i+1}, \ldots, x_m)$$

viewed as a function of ξ, attains a unique minimum $\bar{\xi}$ over X_i, and is monotonically nonincreasing in the interval from x_i to $\bar{\xi}$. Let $\{x^k\}$ be the sequence generated by the block coordinate descent method (2.112). Then, every limit point of $\{x^k\}$ is a stationary point.

Proof: Denote

$$z_i^k = \big(x_1^{k+1}, \ldots, x_i^{k+1}, x_{i+1}^k, \ldots, x_m^k\big).$$

Using the definition (2.112) of the method, we obtain

$$f(x^k) \geq f(z_1^k) \geq f(z_2^k) \geq \cdots \geq f(z_{m-1}^k) \geq f(x^{k+1}), \qquad \forall\ k. \tag{2.113}$$

Let $\bar{x} = \big(\bar{x}_1, \ldots, \bar{x}_m\big)$ be a limit point of the sequence $\{x^k\}$, and note that $\bar{x} \in X$ since X is closed. Equation (2.113) implies that the sequence $\big\{f(x^k)\big\}$ converges to $f(\bar{x})$. We will show that $\bar{x}$ is a stationary point.

Let $\{x^{k_j} \mid j = 0, 1, \ldots\}$ be a subsequence of $\{x^k\}$ that converges to $\bar{x}$. From the definition (2.112) of the algorithm and Eq. (2.113), we have

$$f\big(x^{k_{j+1}}\big) \leq f\big(z_1^{k_j}\big) \leq f\big(x_1, x_2^{k_j}, \ldots, x_m^{k_j}\big), \qquad \forall\ x_1 \in X_1.$$

Taking the limit as j tends to infinity, we obtain

$$f(\bar{x}) \leq f(x_1, \bar{x}_2, \ldots, \bar{x}_m), \qquad \forall\ x_1 \in X_1. \tag{2.114}$$

Using the conditions for optimality over a convex set (Prop. 2.1.2 in Section 2.1), we conclude that

$$\nabla_1 f(\bar{x})'(x_1 - \bar{x}_1) \geq 0, \qquad \forall\ x_1 \in X_1,$$

where $\nabla_i f$ denotes the gradient of f with respect to the component x_i.

The idea of the proof is now to show that $\{z_1^{k_j}\}$ converges to $\bar{x}$ as $j \to \infty$, so that by repeating the preceding argument with $\{z_1^{k_j}\}$ in place of $\{x^{k_j}\}$, we will have

$$\nabla_2 f(\bar{x})'(x_2 - \bar{x}_2) \geq 0, \qquad \forall\ x_2 \in X_2.$$

We can then continue similarly to obtain

$$\nabla_i f(\bar{x})'(x_i - \bar{x}_i) \geq 0, \qquad \forall\ x_i \in X_i,$$

for all $i = 1, \ldots, m$. By adding these inequalities, and using the Cartesian product structure of the set X, it follows that $\nabla f(\bar{x})'(x - \bar{x}) \geq 0$ for all $x \in X$, i.e., $\bar{x}$ is stationary, thereby completing the proof.

To show that $\{z_1^{k_j}\}$ converges to $\bar{x}$ as $j \to \infty$, we assume the contrary, or equivalently that $\{z_1^{k_j} - x^{k_j}\}$ does not converge to zero. Let $\gamma^{k_j} = \|z_1^{k_j} - x^{k_j}\|$. By possibly restricting to a subsequence of $\{k_j\}$, we may assume that there exists some $\overline{\gamma} > 0$ such that $\gamma^{k_j} \geq \overline{\gamma}$ for all j. Let $s_1^{k_j} = \big(z_1^{k_j} - x^{k_j}\big)/\gamma^{k_j}$. Thus, $z_1^{k_j} = x^{k_j} + \gamma^{k_j} s_1^{k_j}$, $\|s_1^{k_j}\| = 1$, and $s_1^{k_j}$ differs from zero only along the first block-component. Notice that $s_1^{k_j}$ belongs to a compact set and therefore has a limit point $\bar{s}_1$. By restricting to a further subsequence of $\{k_j\}$, we assume that $s_1^{k_j}$ converges to $\bar{s}_1$.

Let us fix some $\epsilon \in [0, 1]$. Since $0 \leq \epsilon\overline{\gamma} \leq \gamma^{k_j}$, the vector $x^{k_j} + \epsilon\overline{\gamma} s_1^{k_j}$ lies on the segment joining x^{k_j} and $x^{k_j} + \gamma^{k_j} s_1^{k_j} = z_1^{k_j}$, and belongs to X because X is convex. Using the fact that f is monotonically nonincreasing on the interval from x^{k_j} to $z_1^{k_j}$, we obtain

$$f\big(z_1^{k_j}\big) = f\big(x^{k_j} + \gamma^{k_j} s_1^{k_j}\big) \leq f\big(x^{k_j} + \epsilon\overline{\gamma} s_1^{k_j}\big) \leq f\big(x^{k_j}\big).$$

Since $f(x^k)$ converges to $f(\bar{x})$, Eq. (2.113) shows that $f\big(z_1^{k_j}\big)$ also converges to $f(\bar{x})$. Taking the limit as j tends to infinity, we obtain $f(\bar{x}) \leq f\big(\bar{x} + \epsilon\overline{\gamma}\bar{s}_1\big) \leq f(\bar{x})$. We conclude that $f(\bar{x}) = f\big(\bar{x} + \epsilon\overline{\gamma}\bar{s}_1\big)$, for every $\epsilon \in [0, 1]$. Since $\overline{\gamma}\bar{s}_1 \neq 0$ and by Eq. (2.114), $\bar{x}_1$ attains the minimum of $f(x_1, \bar{x}_2, \ldots, \bar{x}_m)$ over $x_1 \in X_1$, this contradicts the hypothesis that f is uniquely minimized when viewed as a function of the first block-component. This contradiction establishes that $z_1^{k_j}$ converges to $\bar{x}$, which as remarked earlier, shows that $\nabla_2 f(\bar{x})'(x_2 - \bar{x}_2) \geq 0$ for all $x_2 \in X_2$.

By using $\{z_1^{k_j}\}$ in place of $\{x^{k_j}\}$, and $\{z_2^{k_j}\}$ in place of $\{z_1^{k_j}\}$ in the preceding arguments, we can show that $\nabla_3 f(\bar{x})'(x_3 - \bar{x}_3) \geq 0$ for all $x_3 \in X_3$, and similarly $\nabla_i f(\bar{x})'(x_i - \bar{x}_i) \geq 0$ for all $x_i \in X_i$ and i. **Q.E.D.**

Note that the uniqueness of minimum and monotonic nonincrease assumption in Prop. 2.7.1 is satisfied if f is strictly convex in each block-component when all other block-components are held fixed. An alternative assumption under which the conclusion of Prop. 2.7.1 can be shown is that the sets X_i are compact (as well as convex), and that for each i and $x \in X$, the function of the ith block-component ξ

$$f(x_1, \ldots, x_{i-1}, \xi, x_{i+1}, \ldots, x_m)$$

attains a unique minimum over X_i, when all all other block-components are held fixed. The proof is similar (in fact simpler) to the proof of Prop. 2.7.1 [to show that $\{z_1^{k_j}\}$ converges to $\bar{x}$, we note that $\{x_1^{k_j+1}\}$ lies in the compact set X_1, so a limit point, call it $\bar{\xi}$, is a minimizer of $f(x_1, \bar{x}_2, \ldots, \bar{x}_m)$ over $x_1 \in X_1$, and since $\bar{x}_1$ is also a minimizer, it follows that $\bar{\xi} = \bar{x}_1$].

A more general and sometimes useful version of the block coordinate descent method is one where the block-components are iterated in an irregular order instead of a fixed cyclic order. The result of Prop. 2.7.1 can be shown for a method where the order of iteration may be arbitrary as long as there is an integer M such that each block-component is iterated at least once in every group of M contiguous iterations. The proof is similar to the one of Prop. 2.7.1. [Take a limit point of the sequence generated by the algorithm, and a subsequence converging to it, with the properties that 1) every two successive elements of the subsequence are separated by at least M block-component iterations, and 2) every group of M contiguous iterations that starts with an element of the subsequence corresponds to the same order of block-components. Then, proceed as in the proof of Prop. 2.7.1.]

p. 303 (+3) Change "Prop. 3.2.1" to "Prop. 3.2.1, and assume that x^* is a regular point" (The proof of this theorem is correct, but the hypothesis, which was stated correctly in the 1st edition, was inadvertedly corrupted when it was reworded for the 2nd edition. This is also true for the correction in p. 315.)

p. 315 (+14) Change "Prop. 3.3.2" to "Prop. 3.3.2, and assume that x^* is a regular point"

p. 335 (Figure and +5) Change "(2,1)" to "(1,2)"

p. 349 (+7) Replace the portion:
For a feasible x, let $F(x)$ be the set of all feasible directions at x defined by

$$F(x) = \{d \mid d \neq 0, \text{ and for some } \overline{\alpha} > 0,\ g(x + \alpha d) \leq 0 \text{ for all } \alpha \in [0, \overline{\alpha}]\}$$

by the following portion:
For a feasible x, let $F(x)$ be the set consisting of the origin plus all feasible directions at x defined by

$$F(x) = \{d \mid \text{ for some } \overline{\alpha} > 0,\ g(x + \alpha d) \leq 0 \text{ for all } \alpha \in [0, \overline{\alpha}]\}$$

p. 354 (+18) Change "inequality constraints" to "equality constraints"

p. 408 (-2) Change "that" to "than"

p. 418 (+11) Change "(b) Using ..." to "(b) Assume that Q is invertible. Using ..."

p. 423 (Fig. 4.2.8) Change "$\frac{1}{c}(e^{cg} - 1)$" to "$\frac{1}{c}(\mu e^{cg} - 1)$".

p. 423 (Fig. 4.2.8) Change "$-\frac{\mu^2}{2c}$" to "$-\frac{\mu}{c}$".

p. 432 (-12) Change

$$f(x^k + \alpha^k d^k) = \min_{\alpha \in [0,s]} f(x^k + \alpha d^k).$$

to

$$f(x^k + \alpha^k d^k) + cP(x^k + \alpha^k d^k) = \min_{\alpha \in [0,s]} \big\{ f(x^k + \alpha d^k) + cP(x^k + \alpha d^k) \big\}.$$

p. 525 (+8) Change "Exercise 5.4.7" to "Exercise 5.4.6"

p. 525 (+13) Change "Let Assumption 5.4.1 hold." to "Let Assumption 5.4.1 hold and assume that the epigraphs

$$\big\{(x,\gamma) \mid f_1(x) \le \gamma\big\}, \qquad \big\{(x,\gamma) \mid f_2(x) \le \gamma\big\}$$

are closed subsets of $\Re^{n+1}$." (The discussion preceding the proposition assumes the closure of these epigraphs.)

p. 533 (-4) Change "compact" to "closed"

p. 535 (-14) Change "The proof is the same as the one given in Prop. 3.4.3." to "To verify this, take any $u_1, u_2 \in P$, $\alpha \in [0,1]$, and $\epsilon > 0$, and choose $x_1, x_2 \in X$ such that $g(x_i) \le u_i$ and $f(x_i) \le p(u_i) + \epsilon$, $i = 1, 2$. Then, by using the convexity of X, f, and g_j, it is seen that

$$p\big(\alpha u_1 + (1-\alpha)u_2\big) \le \alpha p(u_1) + (1-\alpha)p(u_2) + \epsilon.$$

This implies that $\alpha u_1 + (1-\alpha)u_2 \in P$, so that P is convex, and also, by taking the limit as $\epsilon \to 0$, that p is convex over P."

p. 536 (+13) Change "Exercises 5.4.7 and 5.4.8" to "Exercise 5.4.7"

p. 540 (-5) Change "$\le f(\tilde{x}) + g(\tilde{x})'\mu^* + \epsilon$." to "$\le f(x^*) + g(x^*)'\mu^* + \epsilon$."

p. 549 (-11) Change "closed convex" to "closed"

p. 578 (-8) Change "We eliminate" to "We modify the problem so that the constraint $\sum_{i=1}^m x_{ij} \le T_j y_j$ is replaced by $\sum_{i=1}^m x_{ij} \le T_j$. Then we eliminate"

p. 582 (+2) Change "that" to "than"

p. 585 (5) A correct statement of Exercise 5.5.6 is as follows:
Statement: Let E be a matrix with entries -1, 0, or 1, and at most two nonzero entries in each of its columns. Show that E is totally unimodular if and only if the rows of E can be divided into two subsets such that for each column with two nonzero entries, the following hold: if the two nonzero entries in the column have the same sign, their rows are in different subsets, and if they have the opposite sign, their rows are in the same subset.

p. 609 (+14) Change $\mu\Re^r$ to $\mu \in \Re^r$

p. 615 (+10) Change last "=" to "$\le$"

p. 624 In paragraph starting with "There are several ..." change five instances of "S" to "S^k"

p. 624 (+3 and +9) Delete "$\tilde{q}_k \leq q(\mu)$" from the expression for the set S^k.

p. 635 (-3) Change $\sum_{k=0}^{\infty}(s^k)^2 < \infty$ to $\sum_{k=0}^{\infty}(s^k)^2\|g^k\|^2 < \infty$

p. 649 (+8) Change "is the minimum of the dimensions of the range space of A and the range space" to "is equal to the dimension of the range space of A and is also equal to the dimension of the range space"

p. 651 (-14) Change "if $x_k \leq y_k$," to "if $x_k \leq y_k$ for all k,"

p. 675 (+7) Change "differentiable over C" to "differentiable over $\Re^n$"

p. 676 (+10) Change "differentiable over C" to "differentiable over $\Re^n$"

p. 677 (-9) Change "Convex" to "where α is a positive number. Convex"

p. 704 (+10) Change "al $y \in Y$" to "all $y \in Y$"

p. 705 (+12) Change "$j = 1, \ldots, r$" to "$j \in \{1, \ldots, r\}$"

p. 710 (-2) Change "$< \mu$" to "$\leq \mu$"

p. 725 (-9) Change "... $g(\bar{\alpha} < g(b)$ ..." to "... $g(\bar{\alpha}) < g(b)$ ..."

p. 727 (+6) Change "... $\alpha_{k+1} = \bar{a}_k$ ..." to "... $\alpha_{k+1} = \bar{\alpha}_k$ ...."

p. 764 (-) Change "[RoW97]" to "[RoW98]". Change "1997" to "1998"

p. 765 (-4) Change "Real Analysis" to "Principles of Mathematical Analysis"

Corrections to the 2ND PRINTING

p. 49 (+3 and +4) Change $f(x^k)$ to $\nabla f(x^k)$

p. 100 (+20) Exercise 1.4.4 is flawed and the algorithm needs to be modified. Replace the portion:

Let also d_N^k be the Newton direction $-\big(\nabla^2 f(x^k)\big)^{-1}\nabla f(x^k)$ if $\nabla^2 f(x^k)$ is nonsingular, and be equal to d_S^k otherwise. Consider the method

$$x^{k+1} = x^k + \alpha^k\big((1-\alpha^k)d_S^k + \alpha^k d_N^k\big),$$

where $\alpha^k = \beta^{m^k}$ and m^k is the first nonnegative integer m such that the following three inequalities hold:

$$f(x^k) - f\Big(x^k + \beta^m\big((1-\beta^m)d_S^k + \beta^m d_N\big)\Big) \geq -\sigma\beta^m\nabla f(x^k)'\big((1-\beta^m)d_S^k + \beta^m d_N\big),$$

$$c_1 \min\big\{\nabla f(x^k)'D\nabla f(x^k),\, \|\nabla f(x^k)\|^3\big\} \leq -\nabla f(x^k)'\big((1-\beta^m)d_S^k + \beta^m d_N\big),$$

$$\|(1-\beta^m)d_S^k + \beta^m d_N\| \leq c_2 \max\big\{\|D\nabla f(x^k)\|,\, \|\nabla f(x^k)\|^{1/2}\big\},$$

where β and σ are scalars satisfying $0 < \beta < 1$ and $0 < \sigma < 1/2$, and c_1 and c_2 are scalars satisfying $c_1 < 1$ and $c_2 > 1$.

with the following:

Let also d_N^k be defined as in Exercise 1.4.3. Consider the method

$$x^{k+1} = x^k + \alpha^k\big((1-\alpha^k)d_S^k + \alpha^k d_N^k\big),$$

where $\alpha^k = \beta^{m^k}$ and m^k is the first nonnegative integer m such that

$$f(x^k) - f\big(x^k + \beta^m\big((1-\beta^m)d_S^k + \beta^m d_N\big)\big) \geq -\sigma\beta^m \nabla f(x^k)'\big((1-\beta^m)d_S^k + \beta^m d_N\big),$$

where β and σ are scalars satisfying $0 < \beta < 1$ and $0 < \sigma < 1/2$.

Change also the corresponding part of the hint, i.e., delete "each of the three inequalities is satisfied for m sufficiently large, so" and also "(cf. the last two inequalities)"

p. 134 (-9) Change "be a linear combination" to "either be 0 or be a linear combination"

p. 156 (+14) See correction for page 156 of the 1st printing.

p. 162 (-8) See correction for page 162 of the 1st printing.

p. 214 (+9) Change "closed cone" to "polyhedral cone"

p. 225 (-3) Change x^0 to x^k

p. 226 (+4) Change x_0 to x^k

p. 227 (+7) Change "Zoutendijk's method uses ..." to "Zoutendijk's method rescales d_k so that $x_k + d_k$ is feasible, and uses ..."

p. 273 (+18) See correction for page 268 of the 1st printing.

p. 362 (-3) Change "inequality constraints" to "equality constraints"

p. 428 (+7) Change "some positive" to "any positive"

p. 504 (-1) Change "(Exercise 5.1.3)" to ", see Exercise 5.1.3"

p. 556 (+13) Change "the the" to "that the"

p. 561 (+13) Change "... G is a continuous concave function ..." to "... G is a continuous convex function ..."

p. 576 (-1) Figure 5.5.3 is incorrect. Replace it by the following:

p. 595 (-15) See correction for page 585 of the 1st printing for a correct statement of Exercise 5.5.6.

p. 629 (+4) Change

$$\partial_{\epsilon_1} q(\mu) + \cdots + \partial_{\epsilon_m} q(\mu) \subset \partial_\epsilon q(\mu), \qquad \forall\ \mu \text{ with } q(\mu) > -\infty,$$

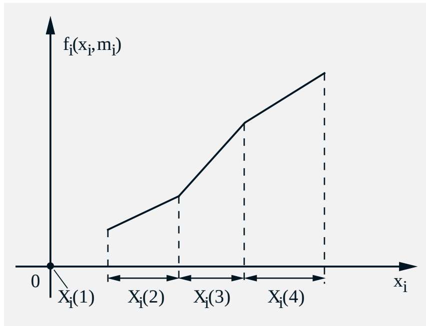

Figure 5.5.3. Production modes and cost function of a discrete resource allocation problem. Here there are four modes of production, $m_i = 1, 2, 3, 4$, and corresponding constraints, $x_i \in X_i(m_i)$. The choice $m_i = 1$ corresponds to no production ($x_i = 0$).

to

$$\partial_{\epsilon_1} q_1(\mu) + \cdots + \partial_{\epsilon_m} q_m(\mu) \subset \partial_\epsilon q(\mu), \qquad \forall\ \mu \text{ with } q(\mu) > -\infty,$$

p. 634 See correction for page 624 of the 1st printing.

p. 634 (-8) Delete "$\tilde{q}_k \leq q(\mu)$" from the expression for the set S^k.

p. 635 (+2) Delete "$\tilde{q}_k \leq q(\mu)$" from the expression for the set S^k.

p. 649 (+10) Change x_i to $\overline{x}_i$ (twice).

p. 660 (-4) Change "... $x = (x_1, \ldots, n)$." to "... $x = (x_1, \ldots, x_n)$."

p. 691 In the figure, change "$f(z)+(z-x)'\nabla f(x)$" to "$f(x)+(z-x)'\nabla f(x)$"

p. 743 (-9) Change "... $g(\bar{\alpha} < g(b)$ ..." to "... $g(\bar{\alpha}) < g(b)$ ..."

p. 745 (+6) Change "... $\alpha_{k+1} = \bar{a}_k$ ..." to "... $\alpha_{k+1} = \bar{\alpha}_k$ ...."

p. 745 (+6) Change "... $g(\bar{b}_k) < g(\alpha_k)$ ..." to "... $g(\bar{b}_k) < g(\bar{\alpha}_k)$ ...."

Corrections to the 3RD PRINTING

p. 116 (-5) Change "$\|x^k - x^*\|$" to "$\big\|x^k - x(\alpha)\big\|$"

p. 184 (-18) Change "6.4.6 in Section 6.4" to "5.4.5 in Section 5.4"

p. 576 See correction for page 576 of the 2nd printing.

p. 649 (-3) Change "$q^* - q(\mu) \leq$ ..." to "$q^* - q(\mu) \geq$ ..."

p. 693 (-9) Change "Prop. Prop." to "Prop."

p. 742 (+19) Change formula

$$z = \frac{3\big(g(a) - g(b)\big)}{b - a} + g'(a) + g'(b),$$

to

$$z = \frac{3\big(g(b) - g(a)\big)}{b - a} + g'(a) + g'(b),$$