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

Note: Many of these corrections have been incorporated in the 2nd Printing of the book. See the end of this file for corrections to the 2nd 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 vctor 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. 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. 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

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

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

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

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

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

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 $x_k \le y_k$," to "if $x_k \le y_k$ 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 "$\le \mu$"

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. 214 (+9) Change “closed cone” to “polyhedral cone”

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. 362 (-3) Change “inequality constraints” to “equality constraints”

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