

## Nonlinear Optimization

by Andrzej Ruszczyński

### Errata

Below is the list of errors in the 1st printing, which were corrected in the 2nd printing. You have the 1st printing if you see “10 9 8 7 6 5 4 3 2 1” on the bottom of the copyright page.

#### First Printing

**page 12, line -3:** Replace  $\|F_{k-1} + G\|$  by  $\|F_{n-1} + G\|$

**page 12, line -2:** Replace “work” by “energy”

**page 12, line -1:** Replace the formula by  $f_2(F) = \sum_{k=1}^{n-1} \|F_k\|^2$

**page 25, line -9:** Replace “2.16” by “2.17”.

**page 29, line -1:** Replace  $\{A^T \lambda : \lambda \in C^\circ\}$  by  $\overline{\{A^T \lambda : \lambda \in C^\circ\}}$ .

**page 30, line 5:** Replace  $\{A^T \lambda : \lambda \in C^\circ\}$  by  $\overline{\{A^T \lambda : \lambda \in C^\circ\}}$ .

**page 35, line -5:** Replace  $\{A^T \lambda : \lambda \in K_2^\circ\}$  by  $\overline{\{A^T \lambda : \lambda \in K_2^\circ\}}$ .

**page 36, lines 10–23:** Replace the text with the following:

$$z = \lim_{j \rightarrow \infty} A^T \lambda^j, \quad \lambda^j \in K_2^\circ.$$

It follows that there exists a sequence  $e^j \rightarrow 0$  such that  $-A^T \lambda^j - e^j \in K_1^\circ$ . Thus, for every  $x \in K_1$  we obtain  $-\langle x, A^T \lambda^j + e^j \rangle \leq 0$ , that is,

$$\langle Ax, \lambda^j \rangle + \langle x, e^j \rangle \geq 0.$$

As  $\lambda^j \in K_2^\circ$ , we also have  $\langle y, \lambda^j \rangle \leq 0$  for all  $y \in K_2$ . Hence

$$\langle Ax - y, \lambda^j \rangle + \langle x, e^j \rangle \geq 0 \quad \text{for all } x \in K_1 \quad \text{and all } y \in K_2.$$

Suppose that  $\|\lambda^j\| \geq \delta > 0$  for an infinite set of indices  $j \in \mathcal{J}$ . Dividing by  $\|\lambda^j\|$  and passing to the limit (over a sub-subsequence, if necessary), we obtain

$$\langle Ax - y, \lambda \rangle \geq 0 \quad \text{for all } x \in K_1 \quad \text{and all } y \in K_2,$$

where  $\lambda$  is an accumulation point of  $\lambda^j/\|\lambda^j\|$  for  $j \in \mathcal{J}$ . Clearly,  $\|\lambda\| = 1$ . Since the set of all  $Ax - y$ , where  $x \in K_1$  and  $y \in K_2$ , contains a neighborhood of 0 by assumption, we obtain a contradiction. Thus we must have  $\lambda^j \rightarrow 0$ , as  $j \rightarrow \infty$ . This yields  $z = 0$  and  $w = 0$ , a contradiction again. Consequently,  $\alpha > 0$ .

Dividing (2.13) by  $\alpha$  and rearranging terms we obtain

$$v = w/\alpha + \lim_{j \rightarrow \infty} A^T(\lambda^j/\alpha), \quad w \in K_1^\circ, \quad \lambda^j \in K_2^\circ.$$

Thus  $v \in K_1^\circ + D^\circ$ , as required.

**page 36, line -2:** Replace “is not needed” by “and the closure operation in  $K^\circ$  are not needed.”

**page 49, line -5:** Replace “=” by “ $\leq$ ”.

**page 60, line -3:** “Lemma 2.36” should be “Lemma 2.70”.

**page 79, line 8:** Replace “=” by “ $\triangleq$ ”. Replace  $\frac{s}{\alpha}$  by  $\frac{s}{\gamma}$ .

**page 85, problem 2.7:** Replace  $K^\circ$  by  $(K_1 \cap \dots \cap K_m)^\circ$ .

**page 97, line -11:** Remove “closed”.

**page 97, line -8:** Replace “Let  $Y_i, i \in I$ , be the given sets and let” by “Let  $Y_i, i \in I$ , be the given sets. Suppose they are compact and let”

**page 98, line 14:** Add “If the sets  $Y_i$  are not compact, we construct their compact subsets satisfying the assumptions. We define a finite set  $S$  by choosing one element in each intersection of  $n + 1$  sets  $Y_i$ , and we replace each  $Y_i$  by  $\text{conv}(Y_i \cap S)$ .”

**page 102, line 6:** Replace  $y_0$  by  $g(x_0)$ .

**page 106, line 16:** “Lemma 3.10” should be “Lemma 3.16”.

**page 106, line -3:** Replace  $\{[g'(x_0)]^T \lambda : \lambda \in [T_{Y_0}(g(x_0))]^\circ\}$  by  $\{[g'(x_0)]^T \lambda : \lambda \in [T_{Y_0}(g(x_0))]^\circ\}$ .

**page 143, line 7:** In formula (3.68) in the middle row replace  $\frac{\tau^2}{2}$  by  $\frac{1}{2}$

**page 167, line 9:** Replace “Theorem 3.33” by “Theorem 3.34”.

**page 167, line 11:** Replace “Theorem 3.46” by “Theorem 3.33”.

**page 167, line -9:** replace “Theorem 3.46” by “Theorem 3.33”.

**page 188, line -8:** In example 4.23, line 3, exchange  $i$  and  $j$ .

**page 197, line 10:** Remove  $q$  from the first sum and  $\hat{\mu}_i$  from the second sum.

**page 198, line 14:** Replace  $\nabla^2 f(\hat{x})$  by  $\nabla_{xx}^2 L(\hat{x}, \hat{\mu})$ .

**page 201, line 4:** add “where  $\hat{z}_i = -g_i(\hat{x})$ .”

**page 205, line -4:** Replace “ $\geq$ ” by “ $\leq$ ”.

**page 206, line 3:** replace  $x_j \in \{0, 1\}$  by  $x_j \in \{-1, 1\}$ .

**page 206, lines 4–5:** Replace “and is not assumed to be positive semidefinite (otherwise 0 is an optimal solution)” by “. Such Problems are used to classify  $n$  objects into two groups, with  $c_{ij}$  representing dissimilarity of objects  $i$  and  $j$ .”

**page 221, line 14:**  $f(x^{k+1}) - f(x^k)$  should be  $f(x^k) - f(x^{k+1})$ .

**page 228, line -1:**  $\|g\|^2$  should be  $\|g\|^4$  (twice).

**page 229, line 7:**  $\|g\|^2$  should be  $\|g\|^4$ .

**page 229, line 15:**  $\|g\|^2$  should be  $\|g\|^4$ .

**page 286, line -4:**  $\frac{1}{M}$  should be  $\frac{2}{M}$ .

**page 329, line -8:** Delete one “ $(d^k, \hat{\lambda}^k, \hat{\mu}^k)$ ”.

**page 332, line -8:** In (6.79) replace  $g(x) = z$  by  $g(x) = -z$ .

**page 332, line -7:** Replace  $g(x_S)$  by  $-g(x_S)$ .

**page 332, line -6:** Replace  $g(x_S)$  by  $-g(x_S)$ .

**page 333, line 1:** Replace  $z = g(\bar{x})$  by  $z = -g(\bar{x})$ .

**page 333, line 8:** Replace  $g_i(x)$  by  $-g_i(x)$ .

**page 334, line 3:** In (6.81) replace  $g(x) = z$  by  $g(x) = -z$ .

**page 344, line 10:** Replace  $\begin{bmatrix} 0 \\ -4 \end{bmatrix}$  by  $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ .

**page 412, line 11:** Replace  $Y = \overline{\text{int } Y}$  by  $Y \subset \overline{\text{int } Y}$ .

**page 442, line -10:** Replace the text by: “Mathematics of Operations Research 31 (2006), 433–452.”

Below is the list of errors in the 1st printing, which remained in the 2nd printing. You have the 2nd printing if you see “10 9 8 7 6 5 4 3 2” on the bottom of the copyright page.

*Second Printing*

**page 67, line -2:** Replace  $\overline{\mathbb{R}}$  by  $\mathbb{R}$ .

**page 104, line -7:** Replace  $\|x_R(\tau) - x_0\|$  by  $x_R(\tau) - x_0$ .

**page 153, line 7:**  $g_i(\hat{x}(b)) < 0$  should be  $g_i(\hat{x}(b)) < b_i$ .

**page 156, line 12:** replace  $x_2 R_3$  by  $x_3 R_3$ .

**page 157, line 15:** remove the second “will”.

**page 230, line -7:** “Theorem 5.11” should be “Theorem 5.7”.

**page 231, line -14:** “Theorem 5.29” should be “Theorem 5.8”.

**page 234, line -1:** There should be minus sign before  $\langle \nabla f(x^k), [\nabla^2 f(x^k)]^{-1} \nabla f(x^k) \rangle$ .

**page 244, line -4 and -1:** There should be no minus sign in the formula for  $\alpha_k$ .

**page 245, line 2:** There should be no minus sign in front of  $\|\nabla f(x^{k-1})\|^2$ .

**page 282, line -3:**  $9x_1 + 16x_2$  should be  $9x_1 + 16|x_2|$ .

**page 282, line -1:**  $\frac{9}{16}$  should be  $\frac{9^2}{16^2}$ .

**page 283, line 5:**  $\frac{9}{16}$  should be  $\frac{9^2}{16^2}$ .

**page 319, line 11:** “Lemma 6.44” should be “Lemma 6.15”.

**page 327, line -1:** The last line should read:  $A_k d^k = -g^0(x^k)$ .

**page 338, line -4:** In (6.85) the second line should be  $ZY \mathbb{1} = \sigma \mathbb{1}$ .

Many thanks to Jonathan Eckstein, Agust Egilsson, John Gubner, Jean-Baptiste Hiriart-Urruty, Robert Jenssen, Dosung Lee, Stephen Robinson, Matthew Saltzman, Hong-Kun Xu, and my Rutgers and Princeton students for pointing to me these inaccuracies.

If you find any errors in the book which are not listed here, please mail me at [rusz@business.rutgers.edu](mailto:rusz@business.rutgers.edu)