Corrections and additions for the book “Perturbation Analysis of Optimization Problems” by J.F. Bonnans and A. Shapiro

Version of Nov. 5, 2017

Some typos in the book that we noticed are of trivial nature and do not need an explanation. There are, however, more subtle corrections that need to be made. There are also simple extensions and additions to the material presented in the book which are worthwhile to mention.

1 Corrections

Section 2.1.2

Minor correction in the proof of theorem 2.17, page 20: one should introduce first $x_1 \in S$ and discuss if $x_0 - x_1$ (instead of x_0) belongs of not to L. The corrected proof is as follows:

Proof. Let $x_1 \in S$, and set $L := \text{cl}\{\text{Sp}(S)\}$. If $x_0 - x_1 \notin L$, then by theorem 2.14, $x_0 - x_1$ and L can be separated, and hence S and $\{x_0\}$ can be separated. If $x_0 - x_1 \in L$, since $x_0 - x_1 \notin \text{ri}(S - x_1)$, then by theorem 2.13, $x_0 - x_1$ can be separated from $(S - x_1)$ by a linear continuous functional over L. This linear functional separates x_0 from S. It can be extended to a continuous linear functional on the space X (see proposition 2.11), and hence the result follows. ∎

Section 2.1.4

In Section 2.1.4, X and X^* are assumed to be paired locally convex topological vector spaces equipped with respective compatible topologies. The closure operation in that section is taken with respect to these topologies. In particular, if X is a nonreflexive Banach space and X^* is its dual, then the standard pair of compatible topologies is the strong (norm) topology in X and weak* topology in X^* (see the discussion after Definition 2.26 on page 25). In that case the closure in the right hand sides of formulas (2.32) and (2.33), while in the space X^*, should be taken in the weak* topology. This may lead to a confusion since in the ”Basic Notation” $\text{cl}(\cdot)$ is defined with respect to the norm (strong) topology. If X is a reflexive Banach space, then one can use strong topologies in X and X^*.

Page 52

In formula (2.115) the inequality $h(\omega) \geq 0$ should be replaced by $h(\omega) \leq 0$. Also, in the following line $x(\omega) > 0$ should be replaced by $x(\omega) < 0$.

Page 66, Proposition 2.90.
For the given definition of the set $\mathcal{T}$ formula (2.171) is in error. Correct definition of the set $\mathcal{T}$ should be

$$\mathcal{T} := \{h \in X : G(x_0) + DG(x_0)h \in K\}. \tag{1.1}$$

In order to see why (2.171) is wrong let us consider the convex set $K \subset Y$. Let $y_0 \in K$ and $T_K(y_0)$ be the corresponding tangent cone. Then $K \in y_0 + T_K(y_0)$ and hence $\operatorname{dist}(d, T_K(y_0)) = 0$ for any $d \in K - y_0$. Formula (2.172) then follows, under Robinson's constraint qualification, by Stability Theorem 2.87. However, if Y is infinite dimensional, it is not true in general that

$$\operatorname{dist}(y, K) = o(\|y - y_0\|), \quad y \in y_0 + T_K(y_0). \tag{1.2}$$

Consider, for example, space $Y := \ell_1$ and the set

$$K := \left\{(y_n) \in \ell_1 : y_1 \geq n y_n^2,\ n = 2, ...\right\}. \tag{1.3}$$

This is a nonempty closed convex set and

$$T_K(0) = \{(y_n) \in \ell_1 : y_1 \geq 0\}. \tag{1.4}$$

We have here that for every n-th coordinate vector e_n and $t > 0$, $te_n \in T_K(0)$, and for $n \geq (2t)^{-1}$,

$$\operatorname{dist}(te_n, K) = \inf_{y \in [0,t]} \{ny^2 + t - y\} = t + (4n)^{-1} \geq t = \|te_n\|. \tag{1.5}$$

Consequently, (1.2) does not hold at $y_0 = 0$.

The assertion of Corollary 2.91 is correct. The fact that, under Robinson's constraint qualification, $T^i_\Phi(x_0) = T_\Phi(x_0)$ and equation (2.173) holds can be easily derived from Stability Theorem 2.87 and the relation that for any *fixed* $d \in T_K(y_0)$ and $y(t) = y_0 + td + o(t)$ it follows that $\operatorname{dist}(y(t), K) = o(t)$, $t > 0$.

Page 68, proof of Prop. 2.95. *Error and correction kindly indicated to us by Nguyen Huy Chieu.* The multifunction $\mathcal{M}(x,t)$ defined at the beginning of the proof is not closed as is wrongly stated. It should be redefined as follows, denoting by $\bar{B}_Y$ the closed unit ball of Y:

$$\mathcal{M}(x,t) := \begin{cases} DG(x_0)x - t((K - G(x_0)) \cap \bar{B}_Y), & \text{if } t \geq 0, \\ \emptyset, & \text{if } t < 0. \end{cases}$$

Page 82. Formula (2.230) of Proposition 2.118 is correct. However, its proof is imprecise. First line of the proof of Proposition 2.118 should be replaced by the following.

It follows from (2.229), applied to f^{**}, that $x^* \in \partial f^{**}(x)$ iff

$$f^{**}(x) = \langle x^*, x\rangle - f^{***}(x^*). \tag{1.6}$$

Since, by the Fenchel-Moreau-Rockafellar Theorem 2.113, $f^{***} = f^*$ the above equation is equivalent to (2.231). Starting from equation (2.231) the proof can be completed as in the book.

Page 215. In the right hand side of formula (3.191), "$FG(x_0)w$" should be replaced by "$DF(x_0)w$".

Page 220. *Error kindly indicated to us by Shaohua Pan, South China University of Technology.* The proof of Prop. 3.88 should be corrected as follows:
Let $x_n := x_0 + t_n h + \frac{1}{2}t_n^2 w_n \in G^{-1}(K)$ be such that $t_n \downarrow 0$ and $t_n w_n \to 0$. A second order expansion of $G(x_n)$ gives

$$G(x_n) = G(x_0) + t_n DG(x_0)h + \tfrac{1}{2}t_n^2 \left(DG(x_0)w_n + D^2G(x_0)(h,h)\right) + o(t_n^2).$$

Since $G(x_n) \in K$, the outer second order regularity of K implies that for some $e_n \to 0$ in Y, $z_n := e_n + DG(x_0)w_n + D^2G(x_0)(h,h)$ belongs to $T_K^2(G(x_0), DG(x_0)h)$. By Prop. 2.97, $Y = DG(x_0)X - T_K(G(x_0))$. Applying Prop. 2.77 to the convex multifunction over X, $\Psi(x) := DG(x_0)x - T_K(G(x_0))$, we deduce the existence of $C > 0$ not depending on n, $w_n' \in X$ and $k_n \in T_K(G(x_0))$, such that $e_n = DG(x_0)w_n' - k_n$, and $\|w_n'\|_X \le C\|e_n\|_Y$. Then $w_n'' := w_n + w_n'$ satisfies.

$$DG(x_0)w_n'' + D^2G(x_0)(h,h) = z_n + k_n.$$

By (3.63) (since the tangent to a convex cone contains this cone), the r.h.s. belongs to $T_K^2(G(x_0), DG(x_0)h)$. Proposition 3.33 implies $w_n'' \in T^2_{G^{-1}(K)}(x_0, h)$. Since $\|w_n'' - w_n\|_X = \|w_n'\|_X \to 0$, the conclusion follows. ∎

Page 242. Line before equation (3.268), "for any $z \in Z$" should be replaced by "for any $z \in \mathcal{C}$".

Page 270. *Errors kindly indicated to us by Alexey F. Izmailov.*
In the first line of (4.21), read $D_x g_i(x_0, u_0)$ (and not $Dg_i(x_0, u_0)$).
In the second line of (4.22), read εd instead of d.

Page 275. Formula (4.39) should be

$$\operatorname{cl}\operatorname{conv}\left(\cup_{x\in\bar{\mathcal{S}}(u_0)} D_u f(x, u_0)\right). \tag{1.7}$$

Under the assumptions of theorem 4.13, the set $\cup_{x\in\bar{\mathcal{S}}(u_0)} D_u f(x, u_0)$ indeed is compact. Therefore, in case the space U is *finite* dimensional, the convex hull of that set is also compact and hence is closed, and hence the topological closure in the above formula can be omitted.

Page 297. *Error kindly pointed to us by Daniel Steck.* For the 'converse' part of Proposition 4.47, we need the additional assumption that the radial cone $\mathcal{R}_K(y_0)$ is closed. This assumption is used when applying (2.32) to derive formula for the normal cone to K_0 (second display, page 298). This converse part is used only in the second paragraph of page 374, where K is a polyhedron, so that the additional assumption is satisfied there.

Page 302. Line before equation (4.131), "$h_n := x_n - x_0$" should be replaced by "$h_n := \kappa_n^{-1}(x_n - x_0)$".

Pages 404. In equation (5.11): read $F(x_0, u_0)$ instead of $D_x F(x_0, u_0)$.

Page 420. *Error kindly indicated to us by Diethard Klatte and Ebrahim Sarabi.* Theorem 5.20 is correct, but the proof of the converse part, starting p. 420 on line 14, must be corrected in the following way. We first recall a classical property of strong regularity, and then state the corrected proof.

A classical property of strong regularity Strong regularity (SR) for the inclusion

$$\varphi(z) + N(z) \ni 0 \tag{1.8}$$

is defined as the local existence of a univalued, Lipschitz inverse for the perturbed linearized system

$$\varphi(z_0) + D\varphi(z_0)(z - z_0) + N(z) \ni \delta. \tag{1.9}$$

We claim that SR holds iff the local existence of a univalued, Lipschitz inverse holds for the perturbed system

$$\varphi(z) + N(z) \ni \delta. \tag{1.10}$$

Indeed, suppose that the SR holds for the linearized system (1.9). It is shown in Theorem 5.13 that this implies that the system (1.10) has a local solution. Proof of this is based on a fixed point theorem (contraction principle). Converse implication of existence of local solution can be proved in a similar way.

As far as Lipschitz continuity is concerned we can proceed as follows. We may by a translation argument assume that $z_0 = 0$ and that $\varphi(z_0) = 0$. We must analyze the relation

$$Az + N(z) \ni \delta, \tag{1.11}$$

where $A := D\varphi(z_0)$. We check first the Lipschitz property. Consider sequences δ_k and δ_k and z_k, z_k', all converging to 0, such that $Az_k + N(z_k) \ni \delta_k$ and $Az_k' +$

$N(z_k') \ni \delta_k'$. We may write these relations as $\varphi(z_k) + N(z_k) \ni \delta_k + \varphi(z_k) - Az_k$, and similarly for z_k', so that by our hypothesis

$$\|z_k' - z_k\| = O(\|\delta_k' - \delta_k' + \varphi(z_k') - Az_k' - \varphi(z_k) - Az_k\|). \tag{1.12}$$

Since the r.h.s. is of order $O(\|\delta_k' - \delta_k'\|) + o(\|z_k' - z_k\|)$, the Lipschitz property follows. Uniqueness of the solution follows whenever δ is close enough to 0.

Proof of the converse part of thm 5.20 (starting p. 420 on line 14). From the uniform second order growth condition, and since $DG(x_0)$ is onto and hence Robinson's constraint qualification holds, theorem 5.17 implies that there exist neighborhoods $\mathcal{V}$ of $0 \in X^* \times Y$ and $\mathcal{V}_X$ of x_0 such that for all $\delta \in \mathcal{V}$, the canonically perturbed problem

$$\operatorname*{Min}_x f(x) - \langle \delta_1, x \rangle; \quad G(x) + \delta_2 \in K. \tag{1.13}$$

has a locally unique solution $x(\delta)$, and since $DG(x_0)$ is onto, there exists a unique $\mu(\delta) \in Y^*$ such that $(x(\delta), \mu(\delta))$ is solution of the canonically perturbed optimality system, i.e.

$$Df(x) + DG(x)^* \mu = \delta_1; \quad G(x) + \delta_2 \in K. \tag{1.14}$$

It remains to show that the mapping $(x(\cdot), \mu(\cdot))$ is Lipschitz continuous. So consider two elements $\hat{\delta}$ and $\tilde{\delta}$ in $\mathcal{V}$, and the associated solutions of (1.14) denoted by $(\hat{x}, \hat{\mu})$ and $(\tilde{x}, \tilde{\mu})$, respectively. Since $DG(x_0)$ is onto, we may apply the stability theorem 2.87 to the relation $G(x) + \delta_2 = 0$. It follows that (reducing the size of the neighborhoods if necessary) there exists $x \in X$ such that

$$G(x) + \hat{\delta}_2 = G(\tilde{x}) + \tilde{\delta}_2; \quad \|\tilde{x} - x\| = O(\|\hat{\delta}_2 - \tilde{\delta}_2\|). \tag{1.15}$$

Let $\delta_1 := Df(x) + DG(x)^* \tilde{\mu}$. Then $(x, \tilde{\mu})$ is a solution of (1.14) for $\delta = (\delta_1, \hat{\delta}_2)$. We have that

$$\delta_1 - \tilde{\delta}_1 = Df(x) - Df(\tilde{x}) + (DG(x) - DG(\tilde{x}))^* \tilde{\mu}.$$

This and (1.15) imply $\|\delta_1 - \tilde{\delta}_1\| = O(\|\hat{\delta} - \tilde{\delta}\|)$. By proposition 4.32 we have that $\|x - \hat{x}\| = O(\|\hat{\delta} - \tilde{\delta}\|)$, and so by (1.15) again, $\|\tilde{x} - \hat{x}\| = O(\|\hat{\delta} - \tilde{\delta}\|)$. Since $DG(x_0)$ is onto, there exists $\varepsilon > 0$ such that, for any x' close enough to x_0, we have that $\|DG(x')^* \mu\| \geq \varepsilon \|\mu\|$. Since $Df(\hat{x}) + G(\hat{x})^* \hat{\mu} = 0$, we deduce that $\|\hat{\mu}\| \leq \varepsilon^{-1} \|Df(\hat{x})\| = O(1)$. Since

$$\begin{aligned} Df(\hat{x}) - Df(\tilde{x}) &= G(\tilde{x})^* \tilde{\mu} - G(\hat{x})^* \hat{\mu} + \tilde{\delta}_1 - \hat{\delta}_1 \\ &= DG(\tilde{x})^* (\tilde{\mu} - \hat{\mu}) + (DG(\tilde{x}) - DG(\hat{x}))^* \hat{\mu} + \tilde{\delta}_1 - \hat{\delta}_1. \end{aligned} \tag{1.16}$$

We deduce that $\|\tilde{\mu} - \hat{\mu}\| = O(\|\tilde{x} - \hat{x}\| + \|\hat{\delta} - \tilde{\delta}\|) = O(\|\hat{\delta} - \tilde{\delta}\|)$. The conclusion follows.

Pages 460-461. In the proof of theorem 5.60, some terms are missing in the optimality system at the end of page 460 and for the quadratic program at the top of page 461. The missing terms are $D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1))$ in the expansion of constraints, and $(D^2_{(x,u)x} G(x_0, u_0)(h_1, u_1))^* \lambda_1$ in the expansion of $D_x L$, as well as the corresponding $\lambda_1 \cdot D^2_{(x,u)x} G(x_0, u_0)((h_1, u_1), h_2)$ in the cost function of the quadratic program. *One should read, starting at the display at the bottom of page 460:*

$$
\begin{aligned}
& D^2_{(x,u)x} L(x_0, \lambda_0, u_0)(h_2, u_2) + D^3_{(x,u)(x,u)x} L(x_0, \lambda_0, u_0)((h_1, u_1), (h_1, u_1)) \\
& \qquad + (D^2_{(x,u)x} G(x_0, u_0)(h_1, u_1))^* \lambda_1 + DG(x_0, u_0)^* \lambda_2 = 0, \\
& D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1)) + Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in \{1, \dots, q\} \cup I^1_+, \\
& D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1)) + Dg_i(x_0, u_0)(h_2, u_2) \leq 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+, \\
& \lambda_{2i} Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+.
\end{aligned}
$$

The above system has a unique solution, since it is the optimality system of the quadratic problem

$$
\begin{array}{ll}
\operatorname{Min}_{h_2} & D^3_{(x,u)(x,u)(x,u)} L(x_0, \lambda_0, u_0)((h_1, u_1), (h_1, u_1), (h_2, u_2)) \\
& \qquad + D^2_{(x,u)(x,u)} L(x_0, \lambda, u_0)((h_2, u_2), (h_2, u_2)) \\
& \qquad + \lambda_1 \cdot D^2_{(x,u)x} G(x_0, u_0)((h_1, u_1), h_2) \\
\text{s.t.} & Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in \{1, \dots, q\} \cup I^1_+, \\
& Dg_i(x_0, u_0)(h_2, u_2) \leq 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+,
\end{array}
$$

whose objective, etc.

Page 476. Second line after equation (5.171), "C^∞-smooth and $\mathcal{G} \pitchfork_x \mathcal{W}_r$" should be replaced by "C^∞-smooth and $\mathcal{G} \pitchfork \mathcal{W}_r$".

Page 516. First line.
The sentence: "If x_0 is a stationary point of (P), then the first order growth condition holds at x_0 iff $C(x_0) = \{0\}$", should be replaced by: "If x_0 is a stationary point of (P) and the extended MF constraint qualification holds, then the first order growth condition holds at x_0 iff $C(x_0) = \{0\}$".

2 Additional material

Page 301. The result of Proposition 4.52 can be extended as follows.

Proposition 2.1 *Suppose that the assumptions of proposition* 4.52 *hold and let* $(\hat{x}(u), \hat{\lambda}(u))$ *be a stationary point of* (P_u) *such that* $\hat{x}(u) \to x_0$ *as* $u \to u_0$. *Then*

$$\|\hat{x}(u) - x_0\| = O\left(\|u - u_0\|^{1/2}\right). \tag{2.17}$$

Proof. It suffices to show that for any sequence u_n converging to u_0 and $x_n := \hat{x}(u_n)$, it follows that $\kappa_n = O\left(\tau_n^{1/2}\right)$, where $\kappa_n := \|x_n - x_0\|$ and $\tau_n := \|u - u_0\|$. Denote also $\lambda_n := \hat{\lambda}(u_n)$. By passing to a subsequence if necessary we can assume that λ_n converges to some $\lambda_0 \in \Lambda(x_0)$.

We have that $f(x_n, u_n) = L(x_n, \lambda_n, u_n)$ and $f(x_0, u_0) = L(x_0, \lambda_0, u_0)$. Because of Robinson's constraint qualification, there exist points x'_n which are feasible for the unperturbed problem (P_{u_0}) and such that $\|x_n - x'_n\| = O(\tau_n)$. Consequently by the second order growth condition we obtain

$$\begin{aligned}
&f(x_n, u_n) - f(x_0, u_0) \\
&\quad = f(x'_n, u_0) - f(x_0, u_0) + f(x_n, u_0) - f(x'_n, u_0) + f(x_n, u_n) - f(x_n, u_0) \\
&\quad \geq c\|x'_n - x_0\|^2 - c_0\tau_n \geq c\kappa_n^2 - c_1\tau_n - c_1\tau_n\kappa_n,
\end{aligned}$$

where c, c_0 and c_1 are some constants with the constant c being positive. Since

$$|L(x_0, \lambda_0, u_n) - L(x_0, \lambda_0, u_0)| = O(\tau_n),$$

it follows that

$$L(x_n, \lambda_n, u_n) - L(x_0, \lambda_0, u_n) \geq c\kappa_n^2 - c_2\tau_n(1 + \kappa_n) \tag{2.18}$$

for some constant c_2. We also have that

$$L(x_n, \lambda_n, u_n) - L(x_n, \lambda_0, u_n) = \langle \lambda_n - \lambda_0, G(x_n, u_n)\rangle$$

and $\langle \lambda_n, G(x_n, u_n)\rangle = 0$. Moreover, since K is generalized polyhedral we can assume by passing to a subsequence if necessary that $\langle \lambda_0, G(x_n, u_n)\rangle = 0$. Together with (2.18) this implies

$$L(x_n, \lambda_0, u_n) - L(x_0, \lambda_0, u_n) \geq c\kappa_n^2 - c_2\tau_n(1 + \kappa_n). \tag{2.19}$$

Consider the mapping $F(z, u) := (D_xL(x, \lambda, u), -G(x, u))$, where $z := (x, \lambda)$, and let $z_n := (x_n, \lambda_n)$ and $z_0 := (x_0, \lambda_0)$. Since the multifunction $\Gamma(z)$ is monotone we obtain by the generalized equations (4.115) that

$$\langle z_n - z_0, F(z_n, u_n) - F(z_0, u_0)\rangle \leq 0. \tag{2.20}$$

On the other hand

$$\begin{aligned}
&\langle z_n - z_0, F(z_n, u_n) - F(z_0, u_0)\rangle \\
&\quad = \langle x_n - x_0, D_xL(x_n, \lambda_n, u_n) - D_xL(x_0, \lambda_0, u_0)\rangle - \langle \lambda_n - \lambda_0, G(x_n, u_n) - G(x_0, u_0)\rangle \\
&\quad = \langle x_n - x_0, D_xL(x_n, \lambda_0, u_n) - D_xL(x_0, \lambda_0, u_0)\rangle \\
&\qquad -\langle \lambda_n - \lambda_0, G(x_n, u_n) - G(x_0, u_0) - D_xG(x_n, u_n)(x_n - x_0)\rangle \\
&\quad = \kappa_n^2 D_{xx}^2 L(x_0, \lambda_0, u_n)(h_n, h_n) + c_3\kappa_n\tau_n + o(\kappa_n^2) \\
&\qquad -\tfrac{1}{2}\langle \lambda_n - \lambda_0, D_{xx}^2 G(x_0, u_0)(h_n, h_n) + O(\tau_n)\rangle \\
&\quad = \kappa_n^2 D_{xx}^2 L(x_0, \lambda_0, u_n)(h_n, h_n) + c_3\kappa_n\tau_n + o(\kappa_n^2) + o(\tau_n)
\end{aligned}$$

for $h_n := (x_n - x_0)/\kappa_n$ and some constant c_3. Together with (2.18) and (2.20) this implies that

$$0 \geq \bar{c}\kappa_n^2 + c_4\kappa_n\tau_n + c_5\tau_n$$

for some constants $\bar{c}$, c_4 and c_5 with the constant $\bar{c}$ being positive. It follows then that $\kappa_n = O\left(\tau_n^{1/2}\right)$, which completes the proof. ∎