Charles Audet • Warren Hare

Derivative-Free and Blackbox Optimization

List of typos

Springer

The most important errors are preceded by a star $\star$ (those on P114 and P123).

Bottom of page 4: Add absolute values around sin

iii. A set of points: such as
$$\underset{x}{\operatorname{argmin}}\{|\sin(x)| : x \in \mathbb{R}, x \in [0,7]\} = \{0, \pi, 2\pi\}.$$

Page 14, Ex 1.3: Replace x by $\bar{x}$ in a) and x, t by $\hat{x}, \hat{t}$ in b)

a) Show that if $\bar{x} \in \operatorname{argmin}(P)$, then $[\bar{x}, f(\bar{x})]^\top$ belongs to $\operatorname{argmin}(Q)$.

b) Show that if $[\hat{x}, \hat{t}]^\top \in \operatorname{argmin}(Q)$, then $\hat{x}$ belongs to $\operatorname{argmin}(P)$.

Page 14, Ex 1.9: Delete extra words

c) Plot the difference between observed and predicted values as a function of $\dot{\gamma}$ using $\eta_0 = 5200, \lambda = 140$ and $\beta = 0.38$, ~~for both $\widehat{g}$ and $\breve{g}$.~~

Page 18, Def 2.2: Change strict equalities to equalities and x to $\bar{x}$

Definition 2.2 (Lipschitz Continuous).
The function $g : \mathbb{R}^n \to \mathbb{R}^m$ is said to be Lipschitz continuous *on the set $X \subseteq \mathbb{R}^n$ if and only if there exists a scalar $K \geqslant 0$ for which*
$$\|g(x) - g(y)\| \leq K\|x - y\| \quad \text{for all } x, y \in X.$$
The scalar K is called ~~the~~ a Lipschitz constant *of g relative to the set X.*
Furthermore, g is said to be locally Lipschitz at $\bar{x} \in \mathbb{R}^n$ (or Lipschitz near $\bar{x} \in \mathbb{R}^n$) if g is Lipschitz on some open ball *centred at $\bar{x}$. That is, there exists $K \geqslant 0$ and $r > 0$ such that*

Top of page 42: Change v_j to v^j

Notice that, for all $k \geq \ell$ there exists a direction $v^k \in \{\pm e_i : i = 1, 2, \ldots, n\}$ such that $x^{k+1} = x^k + \delta^\ell v^k$, and therefore $x^{k+1} = x^\ell + \delta^\ell \sum_{j=\ell}^{k} v^j$.

Bottom of page 42: Function is $\mathscr{C}^0$ not $\mathscr{C}^1$

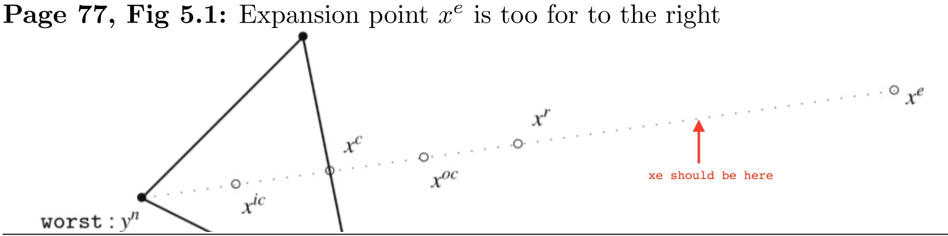
For eventual contradiction, suppose $f'(\hat{x}; d) = -c < 0$. Let $\{x^{k_j}\}$ be a subsequence of unsuccessful iteration that converges to $\hat{x}$. Since $f \in \mathscr{C}^1$, C^0 not C^1
there exists $\bar{t} > 0$ and $\varepsilon > 0$ such that

Page 62, top: Replace C_5^2 by C_2^5
$$\operatorname{prob}(x^2) = \frac{1}{C_5^2} = \frac{2!\,3!}{5!} = \frac{1}{10}.$$

Page 63, Ex 4.5: Typo in vector $[c^1, c^2]^\top$

Example 4.5. Consider $S = \{x \in \mathbb{Z}^2 : 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 4\}$. The natural encoding creates two vectors $c^1 \in \{0,1\}^2$ and $c^2 \in \{0,1\}^3$. The vector $x = [1,4]^\top \in S$ would be encoded as $[c^1, c^2]^\top = \cancel{[0,0,1,1,0]}^\top$ [0, 1, 1, 0, 0] $\in \{0,1\}^5$

Page 77, Fig 5.1: Expansion point x^e is too for to the right



⋆ **Page 114, Definition 6.8:** Inequality should be $\geq$

Definition 6.8 (Clarke Subdifferential).
Let f be a Lipschitz function. For any $x \in \mathbb{R}^n$ define the Clarke subdifferential *as the following set*

$$\partial f(x) := \{v \in \mathbb{R}^n \ : \ f^\circ(x;d) \overset{\geq}{\leq} d^\top v, \ \textit{for all } d \in \mathbb{R}^n\}.$$

⋆ **Page 123:** Misuse of $\check{z}$ and $\hat{z}$.

Lemma 7.3 (Upper Bound on the Mesh Coarseness).
Let $\{\delta_k\}$ be the sequence of mesh size parameters produced by applying the GPS *algorithm to a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ with bounded level sets. Then, there exists ~~a positive~~ an integer $\check{z}$ such that $\delta^k \leq \delta^0(\tau)^{\check{z}}$ for every integer $k \geq 0$.*

The 2 last lines of the proof of Lemma 7.3:

Thus, every δ^k is bounded above by $2\gamma\|G^{-1}\|\tau^{-1}$ and the result follows by selecting the integer $\check{z}$ so that it satisfies $\delta^0(\tau)^{\check{z}} \geq 2\gamma\|G^{-1}\|\tau^{-1}$. $\square$

The first paragraph of the proof of Theorem 7.4:

Proof. Suppose by way of contradiction that there exists an integer $\check{z}$ such that $0 < \delta^0(\tau)^{\check{z}} \leq \delta^k$ for all $k \geq 0$. By Lemma 7.3 and the GPS parameter update rules, we have that for any $k \geq 0$, $\delta^k = \delta^0(\tau)^{r^k}$ where r^k takes its value from the integers of the finite set $\{\check{z}, \check{z}+1, \ldots, \hat{z}\}$, for some $\check{z} \in$ ~~$\mathbb{N}$~~ $\mathbb{Z}$.

Page 132, Ex 7.9: p_i should be replaced by p

$$p = -a + bU, \ \text{for some } a \in \mathbb{Z}, |a| \leq R, b \in \mathbb{Z}, |b| \leq R.$$

Page 134, Project 7b): The matrices should be

$$D = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \end{bmatrix},$$

$$\mathbb{D}^0 = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\mathbb{D}^{k+1} = \mathbb{D}^k, \qquad \text{if iteration } k \text{ succeeds}$$

$$\mathbb{D}^{k+1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \mathbb{D}^k, \qquad \text{if iteration } k \text{ fails.}$$

Page 136, after Def 8.1: A positive spanning set rather than a positive basis

In the GPS algorithm, the mesh size parameter is used to control the poll step, by creating the *poll set* (renamed with a subscript)

$$P^k_{\mathsf{GPS}} = \{x^k + \delta^k d \ : \ d \in \mathbb{D}^k_{\mathsf{GPS}}\},$$

where $\mathbb{D}^k_{\mathsf{GPS}}$ is a positive ~~basis~~ spanning set selected from $\mathbb{D}$ the columns of the matrix

Page 140, first line: Missing noms around d

140 8. MESH ADAPTIVE DIRECT SEARCH

the MADS algorithm, we have $\delta^k \|d\| \leq \Delta^k b$ for all $d \in \mathbb{D}^k_\Delta$. Since $\Delta^k > \delta^k$

Page 141, Th 8.1: Change δ^k to Δ^k

Moreover, $\{\delta^k\}_{k\in K}$ is a subsequence (for some subset of indices K) with $\lim_{k\in K} \delta^k = 0$ if and only if $\{\delta^k\}_{k\in K}$ is a subsequence with $\lim_{k\in K} \Delta^k = 0$.

\Delta

Page 172, in Def 9.8: Remove boldface

DEFINITION 9.8 (Poised for Quadratic Interpolation).
The set $\mathbb{Y} = \{y^0, y^1, \ldots, y^m\} \subset \mathbb{R}^n$ *with* $m = \frac{1}{2}(n+1)(n+2) - 1$, *is* poised for quadratic interpolation *if the system*

0 should not be in boldface

$$\alpha_0 + \alpha^\top y^i + \frac{1}{2}(y^i)^\top H y^i = \mathbf{0}, \quad i = 0, 1, 2, \ldots m$$

Page 180, in Eq (13): Replace Δ^2 by Δ

(9.13)

$$\|\nabla f(y) - \alpha\| \leq \left(K + \frac{K}{6}\sqrt{n}\|\widehat{L}^{-1}\|\right) \Delta^{\cancel{2}\,1}, \quad \text{for all } y \in B_\Delta(y^0).$$

Page 198, in ex 10.7: The variable t is missing

exist $\bar{t} > 0$ such that

$$f(\bar{x} - t\tilde{g}_\Delta(\bar{x})) \leq f(\bar{x}) - \eta t\|\tilde{g}_\Delta(\bar{x})\|^2$$

for all $0 < t < \bar{t}$.

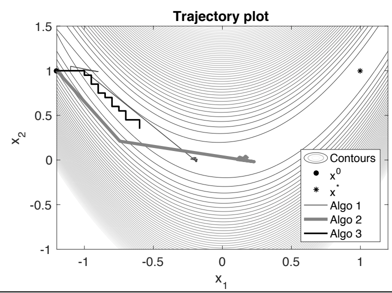
Page 209, Eq (11.9) and Page 215 ex 11.9, 11.10, 11.11: Replace $+$ by $-$

$$(11.9) \qquad \tilde{f}^k(x^k_{\mathtt{tmp}}) \leq \tilde{f}^k(x^k) \overset{-}{\cancel{+}} \frac{1}{2}\mu_C \|\tilde{g}^k\| \min\left\{\frac{\|\tilde{g}^k\|}{\|H^k\|}, \Delta^k\right\},$$

Bottom of page 242: Extra character before the words MADS *with models*

The appendix presents techniques to compare the performance of different algorithms. Figure 13.4 shows the accuracy profile (see Section A.3.4) for the three algorithms CS, MADS, and MADS with models. Each of the 20 test

Page 268: Error in legend



Page 283: Typo in solutions to Ex 4.2b) and 4.4a)

b) $[\cancel{0}, -5/8, 15/4]$. [1/8, -5/8, 14/4]^T

4.4 a) $P_{N=1}(x^2) = 1/5$, $P_{N=2}(x^2) = \cancel{0.13}$, 0.139 $P_{N=5}(x^2) = 1/10$.

Page 300: Typos in index (letter r missing)