PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
S 0002-9939(07)08789-8
Article electronically published on March 1, 2007

THE PROBLEM OF MINIMIZING LOCALLY A C^2 FUNCTIONAL AROUND NON-CRITICAL POINTS IS WELL-POSED

BIAGIO RICCERI

(Communicated by Jonathan M. Borwein)

ABSTRACT. In this paper, we prove the following general result: Let X be a real Hilbert space and $J : X \to \mathbf{R}$ a C^1 functional, with locally Lipschitzian derivative.

Then, for each $x_0 \in X$ with $J'(x_0) \neq 0$, there exists $\delta > 0$ such that, for every $r \in]0, \delta[$, the restriction of J to the sphere $\{x \in X : \|x - x_0\| = r\}$ has a unique global minimum toward which every minimizing sequence strongly converges.

In the sequel, $(X, \langle \cdot, \cdot \rangle)$ is a real Hilbert space. For each $x \in X$, $r > 0$, we set

$$B(x, r) = \{y \in X : \|y - x\| \leq r\}$$

and

$$S(x, r) = \{y \in X : \|y - x\| = r\} \ .$$

Given a functional $J : X \to \mathbf{R}$ and a set $C \subseteq X$, we say that the problem of minimizing J over C is well-posed if the following two conditions hold:

- the restriction of J to C has a unique global minimum, say $\hat{x}$;
- for every sequence $\{x_n\}$ in C such that $\lim_{n\to\infty} J(x_n) = \inf_C J$, one has $\lim_{n\to\infty} \|x_n - \hat{x}\| = 0$.

Assuming $J \in C^1(X)$, our basic question is: under which assumptions is the problem of minimizing J over $S(0, r)$ well posed for each $r > 0$ small enough?

To introduce our result in response to this question, let us make some considerations.

First, consider the case where J is even. Then, for any $r > 0$, $J_{|S(0,r)}$ has either none or at least two global minima. Note that $J'(0) = 0$, since J' is odd.

However, even if $J'(0) \neq 0$, it may occur either that $J_{|S(0,r)}$ has at least two global minima for each $r > 0$, or that $J_{|S(0,r)}$ has no global minima for each $r > 0$. In this connection, consider the following two examples.

Example 1. Take $X = \mathbf{R}^2$ and

$$J(x, y) = x - |y|^q$$

Received by the editors March 22, 2006.
2000 *Mathematics Subject Classification.* Primary 49K40, 90C26, 90C30; Secondary 49J35.
Key words and phrases. Minimization, well-posedness, Hilbert spaces, non-critical points, locally Lipschitzian derivative, saddle points.

where $1 < q < 2$. Note that $J \in C^1(\mathbf{R}^2)$ and $\nabla J(0) \neq 0$. Let $r > 0$. Since

$$\lim_{n\to\infty} n^{q-1}(n - \sqrt{n^2-1}) = 0,$$

for $n \in \mathbf{N}$ large enough, we have

$$J\left(-\sqrt{r^2 - \frac{r^2}{n^2}}, \frac{r}{n}\right) = -\sqrt{r^2 - \frac{r^2}{n^2}} - \left(\frac{r}{n}\right)^q < -r = J(-r,0)\ .$$

Now, observe that $J_{|S(0,r)}$ attains its infimum at some point (x_0, y_0) with $x_0 \leq 0$. The above inequality shows that $x_0 > -r$ (and so $y_0 \neq 0$). Consequently, $(x_0, -y_0)$ is also a global minimum of $J_{|S(0,r)}$.

Example 2. Take $X = l^2$ and

$$J(x) = x_1 - \left(\sum_{n=2}^{\infty} a_n^2 x_n^2\right)^p,$$

where $\frac{1}{2} < p < 1$ and $\{a_n\}$ is a strictly increasing sequence of positive numbers converging to 1. Note that $J \in C^1(l^2)$ and $J'(0) \neq 0$. Fix $r > 0$. Let $\{e_n\}$ be the canonical basis of l^2. Moreover, set

$$I := \{x \in l^2 : x_1 = 0\}$$

and let $A : l^2 \to l^2$ be the operator defined by

$$A(x) = \{a_n x_n\}$$

for all $x \in l^2$. Note that $\|A(e_n)\| = a_n$ and so $\sup_{n\in\mathbf{N}} \|A(e_n)\| = 1$. Note also that $\|A(y)\| < 1$ for all $y \in I \cap S(0,1)$. Further, it is easy to see that

$$S(0,r) = \{-r\sqrt{1-\lambda^2}e_1 + \lambda r y : \lambda \in [0,1], y \in I \cap S(0,1)\}\ .$$

Consequently, we have

$$\inf_{S(0,r)} J = \inf_{y\in I\cap S(0,1)} \inf_{\lambda\in[0,1]} -r\left(\sqrt{1-\lambda^2} + r^{2p-1}\|A(y)\|^{2p}\lambda^{2p}\right)\ .$$

Now, let $\eta : [0,+\infty[\to \mathbf{R}$ be the continuous function defined by

$$\eta(t) = \sup_{\lambda\in[0,1]} (\sqrt{1-\lambda^2} + t\lambda^{2p})$$

for all $t \geq 0$. Since $p < 1$, one readily sees that η is strictly increasing. Hence, we have

$$\inf_{S(0,r)} J = -r \sup_{y\in I\cap S(0,1)} \sup_{\lambda\in[0,1]} \left(\sqrt{1-\lambda^2} + r^{2p-1}\|A(y)\|^{2p}\lambda^{2p}\right)$$
$$= -r \sup_{y\in I\cap S(0,1)} \eta(r^{2p-1}\|A(y)\|^{2p}) = -r\eta(r^{2p-1})\ .$$

But, for every $\lambda \in [0,1]$ and $y \in I \cap S(0,1)$, we have

$$J(-r\sqrt{1-\lambda^2}e_1 + \lambda r y) \geq -r\eta(r^{2p-1}\|A(y)\|^{2p}) > -r\eta(r^{2p-1})$$

and hence $J_{|S(0,r)}$ has no global minima.

Note that, in the above examples, J' is not locally Lipschitzian at 0.

We can now state our main result.

Theorem 1. *Let $J : X \to \mathbf{R}$ be a C^1 functional with locally Lipschitzian derivative.*

Then, for each $x_0 \in X$ with $J'(x_0) \neq 0$, there exists $\delta > 0$ such that, for every $r \in]0, \delta[$, one has

$$\inf_{B(x_0,r)} J = \inf_{S(x_0,r)} J,$$

and the problems of minimizing J over $S(x_0, r)$ and over $B(x_0, r)$ are well-posed.

Proof. Fix $x_0 \in X$ with $J'(x_0) \neq 0$. Also fix $\rho > 0$ so that

$$J'(x) \neq 0$$

for all $x \in B(x_0, \rho)$ and

$$L := \sup_{x,y\in B(x_0,\rho),x\neq y} \frac{\|J'(x) - J'(y)\|}{\|x - y\|} < +\infty \ .$$

For each $\lambda > 0$, $x \in X$, set

$$I_\lambda(x) = \frac{\lambda}{2}\|x - x_0\|^2 + J(x) \ .$$

Let $\lambda \geq L$. For each $x, y \in B(x_0, \rho)$, we have

$$\begin{aligned} \langle I'_\lambda(x) - I'_\lambda(y), x - y\rangle &= \langle \lambda(x - x_0) + J'(x) - \lambda(y - x_0) - J'(y), x - y\rangle \\ &\geq \lambda\|x - y\|^2 - \|J'(x) - J'(y)\|\|x - y\| \geq (\lambda - L)\|x - y\|^2 \ . \end{aligned} \tag{1}$$

From (1), via a classical result ([3, Proposition 25.10]) we then get that the functional I_λ is strictly convex (resp. convex) in $B(x_0, \rho)$ if $\lambda > L$ (resp. $\lambda = L$). Denote by Γ the set of all global minima of the restriction of I_L to $B(x_0, \rho)$ and set

$$\delta = \inf_{x\in\Gamma} \|x - x_0\| \ .$$

Observe that $\delta > 0$. Indeed, if $\delta = 0$, then x_0 would be a local minimum in X for I_L, and so

$$0 = I'_L(x_0) = J'(x_0)$$

against an assumption. Now, fix $r \in]0, \delta[$ and consider the function $\Phi : B(x_0, \rho) \times [L, +\infty[\to \mathbf{R}$ defined by

$$\Phi(x, \lambda) = I_\lambda(x) - \frac{\lambda r^2}{2}$$

for all $(x, \lambda) \in B(x_0, \rho) \times [L, +\infty[$. As we have seen above, $\Phi(\cdot, \lambda)$ is continuous and convex in $B(x_0, \rho)$ for all $\lambda \geq L$, while $\Phi(x, \cdot)$ is continuous and concave for all $x \in B(x_0, \rho)$, with $\lim_{\lambda\to+\infty} \Phi(x_0, \lambda) = -\infty$. So, we can apply to Φ a classical saddle-point theorem [4, Theorem 49.A] which ensures the existence of $(\hat{x}, \hat{\lambda}) \in B(x_0, \rho) \times [L, +\infty[$ such that

$$\begin{aligned} J(\hat{x}) + \frac{\hat{\lambda}}{2}(\|\hat{x} - x_0\|^2 - r^2) &= \inf_{x\in B(x_0,\rho)} \left(J(x) + \frac{\hat{\lambda}}{2}(\|x - x_0\|^2 - r^2) \right) \\ &= J(\hat{x}) + \sup_{\lambda\geq L} \frac{\lambda}{2}(\|\hat{x} - x_0\|^2 - r^2) \ . \end{aligned}$$

Of course, we have $\|\hat{x} - x_0\| \leq r$, since the sup is finite. But, if it were $\|\hat{x} - x_0\| < r$, we would have $\hat{\lambda} = L$. This, in turn, would imply that $\hat{x} \in \Gamma$, against the fact that $r < \delta$. Hence, we have $\|\hat{x} - x_0\| = r$. Consequently

$$J(\hat{x}) + \frac{\hat{\lambda}r^2}{2} = \inf_{x\in B(x_0,\rho)} \left(J(x) + \frac{\hat{\lambda}}{2}\|x - x_0\|^2 \right) \ .$$

From this, we infer that $\hat{\lambda} > L$ (since $r < \delta$), that $\hat{x}$ is a global minimum of the restriction of J to $S(x_0, r)$ and that each global minimum of the restriction of J to $S(x_0, r)$ is a global minimum of the restriction of the functional $I_{\hat{\lambda}}$ to $B(x_0, \rho)$. Since $\hat{\lambda} > L$, this functional is strictly convex, and so $\hat{x}$ is its unique global minimum in $B(x_0, \rho)$ towards which every minimizing sequences weakly converges ([1, p. 3]). In particular, note that if $\{y_n\}$ is a sequence in $B(x_0, \rho)$ such that $\lim_{n\to\infty} J(y_n) = J(\hat{x})$ and $\lim_{n\to\infty} \|y_n - x_0\| = r$, then

$$\lim_{n\to\infty}\left(J(y_n) + \frac{\hat{\lambda}}{2}\|y_n - x_0\|^2\right) = \inf_{x\in B(x_0,\rho)}\left(J(x) + \frac{\hat{\lambda}}{2}\|x - x_0\|^2\right),$$

and so $\{y_n\}$ converges weakly to $\hat{x}$. Since $\lim_{n\to\infty} \|y_n - x_0\| = \|\hat{x} - x_0\|$ and X is a Hilbert space, it follows that $\lim_{n\to\infty} \|y_n - \hat{x}\| = 0$. This shows that, for each $r \in]0, \delta[$, the problem of minimizing J over $S(x_0, r)$ is well posed.

Again fix $r \in]0, \delta[$. Now, let us show that $\inf_{B(x_0,r)} J = \inf_{S(x_0,r)} J$. To this end, for each $t \in [0, r]$, put

$$\varphi(t) = \inf_{S(x_0,t)} J$$

and denote by x_t the unique global minimum of $J_{|S(x_0,t)}$. Clearly, we have

$$\inf_{B(x_0,r)} J = \inf_{[0,r]} \varphi\ .$$

Note also that, by the mean value theorem, J is Lipschitzian in $B(x_0, \rho)$, with Lipschitz constant $L_1 := \|J'(x_0)\| + L\rho$. Fix $t, s \in [0, r]$. We have

$$\varphi(s) - \varphi(t) \le J\left(x_0 + \frac{s}{t}(x_t - x_0)\right) - J(x_t) \le L_1|t - s|$$

as well as

$$\varphi(t) - \varphi(s) \le J\left(x_0 + \frac{t}{s}(x_s - x_0)\right) - J(x_s) \le L_1|t - s|\ .$$

Thus, φ is Lipschitzian and so it attains its infimum in $[0, r]$ at a point $\hat{t}$. In other words, we have

$$\inf_{B(x_0,r)} J = J(x_{\hat{t}})\ .$$

Recalling that $J'(x) \neq 0$ for all $x \in B(x_0, r)$, we then infer that $\hat{t} = r$. So, x_r is also the unique global minimum of $J_{|B(x_0,r)}$. Finally, let $\{y_n\}$ be a sequence in $B(x_0, r)$ such that $\lim_{n\to\infty} J(y_n) = J(x_r)$. By a remark above, to get that $\lim_{n\to\infty} \|y_n - x_r\| = 0$, we have to show that $\lim_{n\to\infty} \|y_n - x_0\| = r$. Argue by contradiction. If it was

$$\liminf_{n\to\infty} \|y_n - x_0\| < r,$$

then, for some $\gamma \in]0, r[$, we would have $\|y_n - x_0\| < \gamma$ for infinitely many n, and so

$$\inf_{B(x_0,r)} J = \inf_{B(x_0,\gamma)} J = J(x_\gamma)$$

against the fact that $J'(x_\gamma) \neq 0$. Thus, the problem of minimizing J over $B(x_0, r)$ is also well posed, and the proof is complete. □

Remark. From the above proof, it is clear the local Lipschitzianity of J' serves only to guarantee that, for some $\rho > 0$ and for each λ large enough, the functional $x \to \frac{\lambda}{2}\|x - x_0\|^2 + J(x)$ is strictly convex in the ball $B(x_0, \rho)$. So, Theorem 1 actually holds for C^1 functionals with this latter property (see [2, pp. 135-136]).

Acknowledgment

The author thanks Professor J. Saint Raymond for useful correspondence.

References

[1] A. L. DONTCHEV and T. ZOLEZZI, *Well-posed optimization problems*, Lecture Notes in Mathematics, 1543, Springer-Verlag, 1993. MR1239439 (95a:49002)
[2] B. MORDUKHOVICH, *Variational analysis and generalized differentiation*, vol. II, Springer-Verlag, 2006. MR2191745
[3] E. ZEIDLER, *Nonlinear functional analysis and its applications*, vol. II/B, Springer-Verlag, 1985. MR1033498 (91b:47002)
[4] E. ZEIDLER, *Nonlinear functional analysis and its applications*, vol. III, Springer-Verlag, 1985. MR0768749 (90b:49005)

Department of Mathematics, University of Catania, Viale A. Doria 6, 95125 Catania, Italy
E-mail address: ricceri@dmi.unict.it