Revista de la
Unión Matemática Argentina
Volumen 30, 1982-83

THE PLENARY HULL OF THE GENERALIZED JACOBIAN MATRIX AND THE INVERSE FUNCTION THEOREM IN SUBDIFFERENTIAL CALCULUS

Telma Caputti

INTRODUCTION.

Some of the most important recent advances in optimatization have come about as a result of a systematic replacement of smoothness assumptions by convexity. This is exemplified by the work of Rockafellar [2]. It is natural to ask whether analogous results can be proven without either smoothness or convexity. A general theory of necessary conditions for such problems has been obtained [3]. The conditions are expressed, in part, by means of generalized gradients.

The classical inverse function theorem gives conditions under which a C^r function admits (locally) a C^r inverse. The purpose of this article is to give conditions under which a lipschitzian (not necessarily differentiable) function admits (locally) a lipschitzian inverse by means the characterization of the plenary hull of the generalized jacobian matrix.

1. LOCALLY LIPSCHITZ FUNCTIONS.

Let $F: R^n \longrightarrow R$ be locally Lipschitz on B a bounded subset of R^n. That is, for each bounded subset B of R^n there exists a constant K such that

$$|f(x_1)-f(x_2)| \leqslant K\|x_1-x_2\| \quad \text{for all points } x_1 \text{ and } x_2 \text{ of B.}$$

It is known [1] that such function has at almost all points x a derivative (gradient), which we denote $\nabla f(x)$. It is easily verified that the function ∇f is bounded on bounded subsets of its domain of definition.

Let now F be locally Lipschitz on O a nomempty open subset of R^n and taking values in R^m. One could be tempted to define the generalized derivative of $F = (f_1,\dots,f_m)^t$ at $x_0 \in O$ by simply considering $[\partial f_1(x_0),\dots,\partial f_m(x_0)]^t$ where this set consists of matrices whose ith row belongs to $\partial f_i(x_0)$.

The usual mxn jacobian matrix of partial derivatives, when it exists, is denoted $JF(x)$. We topologize the vector space of mxn matrices with the norm

$$\|M\| = \max|m_{ij}| \text{ where } M = (m_{ij}) \quad , \quad 1 \leqslant i \leqslant m \quad , \quad 1 \leqslant j \leqslant n$$

A mathematical tool is what Clarke, F.H. called the generalized jacobian matrix defined in the following way:

DEFINITION 1.1. The generalized jacobian matrix of F at $x_0 \in O$, denoted by $\widetilde{J}F(x_0)$, is the *convex hull* of all matrices M of the form $M = \underset{n\to\infty}{\text{Lím}}\ JF(x_n)$ where x_n converges to x_0 in domF'.

In this definition, domF' denotes the subset of full measure of O where F is differentiable.

In doing so, $\widetilde{J}F(x_0)$ is nonempty compact convex subset of the vector space of mxn matrices, which is reduced to $\{JF(x_0)\}$ whenever F is strictly differentiable at x_0.

DEFINITION 1.2. $\widetilde{J}F(x_0)$ is said to be of *maximal rank* if every M in $\widetilde{J}F(x_0)$ is of maximal rank.

2. PLENARY HULL OF THE GENERALIZED JACOBIAN MATRIX.

Let us denote $\langle\!\langle\,,\rangle\!\rangle$ the inner product on the vector space of mxn matrices defined by $\langle\!\langle M,U\rangle\!\rangle$ = Trace of $M \circ U^t$; it comes from Definition 1.1. that for all $U \in R^{m\times n}$:

$$\max_{M\in\widetilde{J}F(x_0)} \langle\!\langle M,U\rangle\!\rangle = \underset{\substack{x\to x_0\\ x\in \mathrm{dom}F'}}{\text{Lím sup}} \langle\!\langle JF(x),U\rangle\!\rangle \tag{2.1}$$

We turn our attention to those $U \in R^{m\times n}$ of the form $u \otimes v$: $x \longrightarrow \langle u,x\rangle v$, where $u \in R^n$ and $v \in R^m$. In such a way, $\langle\!\langle M, u\otimes v\rangle\!\rangle$ reduces to $\langle Mu,v\rangle$ and (2.1) can be rephrased as:

$$\max_{M\in\widetilde{J}F(x_0)} \langle Mu,v\rangle = \underset{\substack{x\to x_0\\ x\in \mathrm{dom}F'}}{\text{Lím sup}} \langle JF(x)u,v\rangle \tag{2.2}$$

We can use results on chain rules so that the left-hand of (2.2) appears as the generalized gradient of a particular real-valued function. Given $v \in R^m$, the *generalized gradient* of F_v: $x \longrightarrow \langle F(x),v\rangle$ at x_0 can be exactly described as: $\partial F_v(x_0) = \widetilde{J}^tF(x_0)v$, ([6]).

Therefore, for all $u \in R^n$, we have that:

$$\max_{M\in \tilde{J}F(x_0)} \langle u, M^t v\rangle = F_v^o(x_0;u) \qquad (2.3)$$

Although $\tilde{J}F(x_0)$ is convex and compact, one generally cannot separate an M_0 from $\tilde{J}F(x_0)$ by using only linear mappings (on $R^{m\times n}$) of the form $u\otimes v$, $u \in R^n$, $v \in R^m$. This state of affairs led Sweetser [5] to introduce the following definition:

- a subset $A \subset R^{m\times n}$ is *plenary* if and only if it includes every M in $R^{m\times n}$ satisfying $Mu \in Au$ for all $u \in R^n$ -

Since the intersection of plenary sets is plenary, Sweetser defined the *plenary hull of* A, denoted *plen* A, as

- a smallest plenary set containing A -

Namely, when $\min(m,n) > 1$, plen $\tilde{J}F(x_0)$ is a convex compact (plenary) set of matrices containing $\tilde{J}F(x_0)$.

Since $\tilde{J}F(x_0)u = [\text{plen } \tilde{J}F(x_0)]u$ for all $u \in R^n$, Hiriart-Urruty [8], formulated the following theorem:

THEOREM 2.1. *Let* $u \in R^n$ *and* $v \in R^m$ *Then,*

$$\max_{M\in \text{plen}\tilde{J}F(x_0)} \langle Mu, v\rangle = F^o(x_0;u,v) \qquad (2.4)$$

In another setting, $M \in \text{plen } \tilde{J}F(x_0)$ if and only if

$$\langle Mu, v\rangle \leqslant F^o(x_0;u,v) \quad \text{for all} \quad (u,v) \in R^n \times R^m.$$

To summarize, let's say that plen $\tilde{J}F(x_0)$ is the convex compact (plenary) set of matrices satisfying

$[\text{plen } \tilde{J}F(x_0)]u = \tilde{J}F(x_0)u$ for all $u \in R^n$. When $F = (f_1,\dots,f_m)^t$ we have that:

$$\tilde{J}F(x_0) \subset \text{plen } \tilde{J}F(x_0) \subset [\partial f_1(x_0),\dots,\partial f_m(x_0)]^t$$

The set $[\partial f_1(x_0),\dots,\partial f_m(x_0)]^t$ is obviously convex, compact and plenary. It actually yields the same image set as $\tilde{J}F(x_0)$ does when the considered vectors u are the elements e_i of the canonical basis in R^n. In other words,

$$\{x_i^*, [x_1^*,\dots,x_i^*,\dots,x_m^*]^t \in \tilde{J}F(x_0)\} = \partial f_i(x_0) \text{ [6]}.$$

3. THE PLENARY HULL OF $\tilde{J}F(x_0)$ AND THE INVERSE FUNCTION THEOREM.

THEOREM 3.1. *If every matrix* M *in* plen $\tilde{J}F(x_0)$ *is of maximal rank, then exist neighborhoods* U *and* V *of* x_0 *and* $F(x_0)$ *respectively, and a lipschitzian function* $G: V \longrightarrow R^n$ *such that:*

for all $u \in U$ *is* $F(u) = v$ *if and only if for all* $v \in V$ *is* $G(v) = u$.

When F is C^1, $\tilde{J}F(x_0)$ reduces to $JF(x_0)$ and the function G above is necessarily C^1 as well. Thus we recover the classical theorem.

REMARK 1. This theorem remains true (without modifications in proof) if we impose the maximality of a rank for all $M \in \tilde{J}_\Lambda F(x_0)$ where $\Lambda \subset \mathrm{dom}F'$ be such that its complementary set in O is of null measure and let $\tilde{J}_\Lambda F(x_0)$ be defined as in (1.1) except that in this definition the points x_n are constrained in Λ.

REMARK 2. Due to the definitions themselves, we have that

$$\partial_\Lambda F_v(x_0) = \tilde{J}^t_\Lambda F(x_0)v \qquad ([6])$$

It is known that the generalized gradient of real-valued functions is blind to sets of null measure [3, Proposition 1.11]. The desire to make the generalized derivative blind to sets of null measure led Porciau, B.H. [7] to alter Clarke's original definition by considering the Lebesgue set the LebF' of F', instead of domF', in the definition of $\tilde{J}F(x_0)$ but, since F' is locally in $L^\infty(0,R^m)$, almost every x in domF' belongs to LebF'.

REMARK 3. Let Λ and $\tilde{J}_\Lambda F(x_0)$ be as in Remark 1. Then

$$\text{plen } \tilde{J}_\Lambda F(x_0) = \text{plen } \tilde{J}F(x_0)$$

So, plen $\tilde{J}F(x_0)$ is blind to sets of measure zero.

Proof of the theorem 3.1.

LEMMA 1. *An exact chain rule in finite-dimensional case* [6].

Let F *be a locally Lipschitz function and* g *be continuously differentiable. Then*

$$\partial(g \circ F) = \tilde{J}^t F(x_0)\ \nabla g(F(x_0)) \textit{ where } F: R^n \longrightarrow R^m \textit{ and } g: R^m \longrightarrow R\ .$$

LEMMA 2. *Let* β *be a positive number. Then for all* x *sufficiently near* x_0, plen $\tilde{J}F(x) \subset [\text{plen } \tilde{J}F(x_0) + \beta M(0,1)]$ *where* M(0,1) *denotes the unit ball in the vector space of* mxn *matrices.*

This is a direct consequence of definition of the plenary hull of the generalized jacobian matrix.

LEMMA 3. *There are positive numbers* r *and* λ *with the following property: given any unit vector* v *in* R^n, *there is a unit vector* u *in* R^n *such that, whenever* x *lies in* $(x_0 + rB)$ *and* $M \in \text{plen}\, \tilde{J}F(x_0)$, *then* $\langle Mv,u\rangle \geqslant \lambda$ *for all* M, *where* B *denotes the open unit ball in* R^n.

Proof. Let Σ_1 denote the unit sphere in R^n. The subset $\text{plen}\, \tilde{J}F(x_0)\, \Sigma_1$ of R^n is compact and does not contain $\underline{0}$ since $\text{plen}\, \tilde{J}F(x_0)$ is of maximal rank.

Hence for some $\lambda > 0$, $\text{plen}\, \tilde{J}F(x_0)\, \Sigma_1$ is distance at least 2λ from $\underline{0}$. For positive β sufficiently small, $[\text{plen}\, \tilde{J}F(x_0) + \beta M(0,1)]\, \Sigma_1$ is distance at least λ from $\underline{0}$.

By Lemma 2, it follows that for some positive r,

$$x \in x_0 + rB \Rightarrow \text{plen}\, \tilde{J}F(x_0) \subset [\text{plen}\, \tilde{J}F(x_0) + \beta M(0,1)]$$

We may suppose r chosen so that F satisfies Lipschitz condition on $(x_0 + r\bar{B})$.

Now let any unit vector v be given. It follows from above that the convex set $[\text{plen}\, \tilde{J}F(x_0) + \beta M]\, v = [\tilde{J}F(x_0) + \beta M]\, v$, for all v in R^n, is distance at least λ from $\underline{0}$. By the usual separation theorem for convex sets, there is a unit vector u such that: $\langle u, Mv\rangle \geqslant \lambda$ for all $M \in \text{plen}\, \tilde{J}F(x_0)$.

LEMMA 4. *If* x_1 *and* x_2 *lie in* $(x_0 + r\bar{B})$, *then*

$$\|F(x_1) - F(x_2)\| \geqslant \lambda \|x_1 - x_2\|$$

Proof. We may suppose $x_1 \neq x_2$ and by the continuity of F that $x_1, x_2 \in (x_0 + rB)$.

$$\text{Set } v = \frac{x_2 - x_1}{\|x_2 - x_1\|}, \quad \alpha = \|x_2 - x_1\| \text{ so that } x_2 = x_1 + \alpha v.$$

Let π be the plane perpendicular to v and passing through x_1. The set P of points x in $(x_0 + rB)$ where F' fails to exist is of measure zero, and hence by Fubini's theorem, for almost every x in π, the ray x+tv, $t \geqslant 0$ meets P in a set of null one-dimensional measure. Choose an x with the above property and sufficiently close to x_1 so that x+tv lies in $(x_0 + rB)$ for every t in $[0,\alpha]$. Then the function $t \longrightarrow F(x+tv)$ is Lipschitzian for t in $[0,\alpha]$ and has a.e. on this interval the derivative JF(x+tv) v. Thus

$$F(x+\alpha v) - F(x) = \int_0^\alpha JF(x+tv)v\,dt$$

Let u be as in Lemma 3. We deduce,

$$\langle u, (F((x+\alpha v) - F(x))\rangle = \langle u, \int_0^\alpha JF(x+tv)v\; dt\rangle \geqslant \int_0^\alpha \lambda\; dt = \lambda.\alpha$$

Recalling the definition of α, we arrive at:

$$\|F(x+\alpha v) - F(x)\| \geqslant \lambda \|x_2 - x_1\|$$

This may be done for x arbitrarily close to x_1. Since F is continuous, the lemma ensues.

LEMMA 5. $F(x_0 + rB)$ *contains* $F(x_0) + (r\,\lambda/2)B$.

Proof. Let y be any point in $F(x_0) + (r\,\lambda/2)B$, and let the minimum of $\|y - F(x)\|^2$ over $(x_0 + r\bar{B})$ be attained at x. We claim x belongs to $(x_0 + rB)$.

$$\begin{aligned} r\,\lambda/2 > \|y - F(x_0)\| &\geqslant \|F(x) - F(x_0)\| - \|y - F(x)\| \geqslant \\ &\geqslant \lambda\|x - x_0\| - \|y - F(x)\| \geqslant \\ &\geqslant \lambda r - \|y - F(x_0)\| > \lambda r - r\,\lambda/2 = r\,\lambda/2 \end{aligned}$$

which is a contradiction. Thus x yields a local minimum for the function $\|y - F(x)\|^2$, and consequently [3, Corollary 1.10],

$$\underline{0} \in \partial\|y - F(x)\|^2$$

We now use Lemma 1 to conclude that $\underline{0}$ belongs to the set

$$\tilde{J}^t F(x)(y - F(x)) \quad ([6])$$

that coincides with $[\text{plen } \tilde{J}^t F(x)](y - F(x))$ for all vector in R^m by Theorem 2.1. But Lemma 3 implies that every matrix in plen $\tilde{J}F(x)$ is non singular, hence the above is possible only if $F(x) = y$.

We now set $V = F(x_0) + (r\,\lambda/2)B$, and we define G on V as follows:

$G(v)$ is the unique x in $(x_0 + rB)$ such that $F(x) = v$. We choose U as any neighborhood of x_0 satisfying $F(U) \supset V$. The theorem is now seen to follow, since Lemma 4 implies that G is Lipschitz with constant λ^{-1}.

REFERENCES

[1] STEIN,E.M., *Singular integrals and differentiability properties of functions*, Princeton Academic Press. Princeton University. New Jersey - 1970.

[2] ROCKAFELLAR,R.T., *Convex Analysis*, Princeton Academic Press. Princeton University. New Jersey - 1970.

[3] CLARKE,F.H., *Generalized gradients and applications*, Transactions of the American Mathematical Society. Vol.205 (1975). Pag.247-262.

[4] HIRIART-URRUTY,J.B., *Contributions à la programmation mathématique: cas déterministe et stochastique*, Thèse de Doctorat en Sciences Mathématiques. Université de Clermont-Ferrand II - 1977.

[5] SWEETSER,TH., *A minimal set-valued strong derivative for vector-valued Lipschitz functions*, Journal of Optimization Theory and Applications. Vol.23 (1977). Pag.549-562.

[6] HIRIART-URRUTY,J.B., *News concepts in nondifferentiable programming*, Journées d'Analyse Non Convexe. Université de Pau. Mai 1977. Bulletin de la Société Mathématique de France. Mémoire N°60 (1979). Pag.57-85.

[7] POURCIAU,B.H., *Analysis and optimization of Lipschitz continuous mappings*, Journal of Optimization Theory and Applications. Vol. 22 (1977). Pag.311-351.

[8] HIRIART-URRUTY,J.B., *Analyse Fonctionnelle - Existence et caractérisation de différentielles généralisées d'applications localment lipschitziennes d' un space de Banach séparable dans un space de Banach réflexif séparable*, C.R.Acad. Sc. Paris, t.209 - 1980.

Departamento de Matemática
Facultad de Ciencias Exactas
Universidad de Buenos Aires.

Recibido en setiembre de 1981.