A Hamiltonian approach to implicit systems and applications in optimization

Dan TIBA

"SIMION STOILOW" INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY

dan.tiba@imar.ro

CONTENTS

EXAMPLES IN DIMENSION THREE : SURFACES

The question of the parametrization of an implicitly defined manifold.

- The hypothesis $\nabla f(x_0, y_0, z_0) \neq 0$ is specified in the form $f_x(x_0, y_0, z_0) \neq 0$.

- We associate to the equation $f(x, y, z) = 0$, $f(x_0, y_0, z_0) = 0$, two iterated Hamiltonian systems:

$$
\begin{aligned}
x' &= -f_y(x, y, z), \qquad && t \in I_1, \\
y' &= f_x(x, y, z), && t \in I_1, \\
z' &= 0, && t \in I_1, \\
x(0) &= x_0, \quad y(0) = y_0, \quad z(0) = z_0;
\end{aligned}
$$

EXAMPLES IN DIMENSION THREE : SURFACES

- The second system is:

$$
\begin{aligned}
\dot{\varphi} &= -f_z(\varphi,\psi,\xi), && s \in I_2(t),\\
\dot{\psi} &= 0, && s \in I_2(t),\\
\dot{\xi} &= f_x(\varphi,\psi,\xi), && s \in I_2(t),\\
\varphi(0) &= x(t),\ \ \psi(0) = y(t),\ \ \xi(0) = z(t).
\end{aligned}
$$

Comments: dimension two, PDE's and ODE's, numerical solution.

Domain of existence I_2 and its evaluation.

EXAMPLES IN DIMENSION THREE : SURFACES

Proposition

We have:

$$f(\varphi(t,s),\psi(t,s),\xi(t,s)) = 0,\ \forall (t,s) \in I_1 \times I_2$$

In the next proposition, under more regularity assumptions, we obtain a **local** parametrization:

Proposition

If the domain $I_1 \times I_2$ is sufficiently small around the origin and $f \in C^2(\Omega)$, then the mapping $(\varphi,\psi,\xi) : I_1 \times I_2 \to R^3$ is regular and one-to-one on its image.

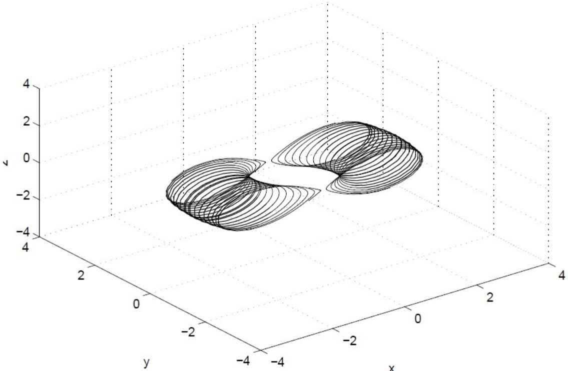
Comparison local, maximal and implicit solutions. Domain of existence.

EXAMPLES IN DIMENSION THREE : SURFACES

- The next two examples are solved with MatLab.

$$3)\ f(x,y,z) = (x^2+y^2+z^2+R^2-r^2)^2 - 4R^2(x^2+y^2)$$

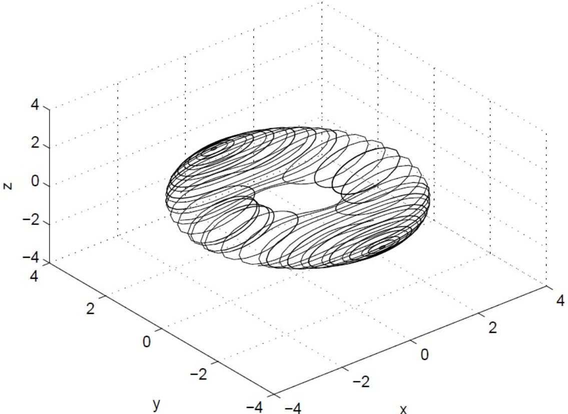
$R=2,\ r=1,\ (x_0,y_0,z_0)=(1,0,0)$



EXAMPLES IN DIMENSION THREE : SURFACES

$$4)\ f(x,y,z) = (x^2+y^2+z^2+R^2-r^2)^2 - 4R^2(x^2+y^2)$$

$R=2,\ r=1,\ (x_0,y_0,z_0)=(3,0,0)$



Arbitrary Dimension

In the Euclidean space R^d, $d \in N$, we consider the general implicit functions system:

$$F_1(x_1, \ldots, x_d) = 0,$$

$$F_2(x_1, \ldots, x_d) = 0,$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$F_l(x_1, \ldots, x_d) = 0,$$

where $l \in N$, $1 \leq l \leq d - 1$ and $F_j \in C^1(\Omega)$, $F_j(x^0) = 0,\ j = \overline{1, l}$, $x^0 \in \Omega \subset R^d$ bounded domain, given.

Arbitrary Dimension

To fix ideas, we assume

$$\frac{D(F_1, F_2, \ldots, F_l)}{D(x_1, x_2, \ldots, x_l)} \neq 0 \quad \text{in } x^0 = (x_1^0, x_2^0, \ldots, x_d^0),$$

which remains valid on a neighbourhood $V \in \mathcal{V}(x^0)$, $V \subset \Omega$, under the $C^1(\Omega)$ assumption on $F_j(\cdot)$, $j = \overline{1, l}$.

We denote by $A(x)$ the corresponding $l \times l$ matrix and the vectors $\nabla F_1(x)$, ..., $\nabla F_l(x)$ are independent, for $x \in V$.

Arbitrary Dimension

We introduce on V the linear systems of equations with unknowns $v(x) \in R^d$, $x \in V$:

$$v(x) \cdot \nabla F_j(x) = 0, \quad j = \overline{1, l}.$$

We shall use $d - l$ solutions obtained by fixing the last $d - l$ components of the vector $v(x) \in R^d$ to be the rows of the identity matrix in R^{d-l}, multiplied by $\det A(x)$. Then, the first l components are uniquely determined, by inverting $A(x)$.

In this way, the obtained $d - l$ solutions, denoted by $v_1(x), \ldots$, $v_{d-l}(x) \in R^d$ are linear independent, for any $x \in V$. Moreover, these vector fields are continuous in V.

Arbitrary Dimension

We introduce now $d - l$ nonlinear systems of first order partial differential equations associated to the vector fields $(v_j(x))_{j=\overline{1,d-l}}$, $x \in V \subset \Omega$.

We use here the order $v_1, v_2, \ldots, v_{d-l}$ to fix ideas. Moreover, we denote the sequence of independent variables by $t_1, t_2, \ldots, t_{d-l}$.

These systems have an iterated character in the sense that the solution of one of them is used as initial condition in the next one. Consequently, the independent variables in the "previous" systems enter as parameters in the next system.

Arbitrary Dimension

$$\frac{\partial y_1(t_1)}{\partial t_1} = v_1(y_1(t_1)), \quad t_1 \in I_1 \subset R,$$

$$y_1(0) = x^0;$$

$$\frac{\partial y_2(t_1, t_2)}{\partial t_2} = v_2(y_2(t_1, t_2)), \quad t_2 \in I_2(t_1) \subset R,$$

$$y_2(t_1, 0) = y_1(t_1);$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\frac{\partial y_{d-l}(t_1, t_2, \dots, t_{d-l})}{\partial t_{d-l}} = v_{d-l}(y_{d-l}(t_1, t_2, \dots, t_{d-l})),$$

$$t_{d-l} \in I_{d-l}(t_1, \dots, t_{d-l-1}),$$

$$y_{d-l}(t_1, \dots, t_{d-l-1}, 0) = y_{d-l-1}(t_1, t_2, \dots, t_{d-l-1}).$$

Arbitrary Dimension

Here, the notations $I_1, I_2(t_1), \ldots, I_{d-l}(t_1, \ldots, t_{d-l-1})$ are $d-l$ real intervals, containing 0 in interior and depending, in principle, on the "previous" parameters.

Due to their simple structure, we stress that each equation may be interpreted as an ordinary differential system with parameters, although partial differential notations are used.

The existence of the solutions $y_1, y_2, \ldots, y_{d-l}$ follows by the Peano theorem due to the continuity of the vector fields $(v_j)_{j=\overline{1,d-l}}$ on V.

Arbitrary Dimension

Proposition

a) There are closed intervals $I_j \subset R$, $0 \in \mathrm{int} I_j$, independent of the parameters, such that $I_j \subset I_j(t_1, t_2, \ldots, t_{j-1})$, $j = \overline{1, d-l}$.
b) For any fixed values of the corresponding parameters, the above differential systems have solutions $y_j \in C^1(I_j)$, $j = \overline{1, d-l}$. Moreover, $y_j, j = \overline{1, d-l}$ are continuously differentiable in the origin and we have:

$$\frac{\partial y_{d-l}}{\partial t_k}(0, \ldots, 0) = v_k(x^0), \quad k = \overline{1, d-l}.$$

Proposition

For every $k = \overline{1, l}$, $j = \overline{1, d-l}$, we have

$$F_k(y_j(t_1, t_2, \ldots, t_j)) = 0, \quad \forall\, (t_1, t_2, \ldots, t_j) \in I_1 \times I_2 \times \ldots \times I_j.$$

Arbitrary Dimension

Proposition

If $F_k \in C^2(\Omega)$, $k = \overline{1, l}$, and I_j are sufficiently small, $j = \overline{1, d-l}$, then the mapping $y_{d-l} : I_1 \times I_2 \times \ldots \times I_{d-l} \to R^d$ is regular and one-to-one on its image.

The local solution of the initial system is a $d - l$ dimensional manifold around x^0 and $y_{d-l}(t_1, t_2, \ldots, t_{d-l})$ is a local parametrization of this manifold on $I_1 \times I_2 \times \ldots \times I_{d-l}$. Choose the last $d - l$ components of the solutions $v_j(x) \in R^d$ as the rows of the identity matrix in R^{d-l}. We obtain :

Proposition

The last $d - l$ components of y_{d-l} have the form $(t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_j + x^0_{l+j}, x^0_{l+j+1}, \ldots, t_{d-l} + x^0_d)$, that is the first l components of y_{d-l} give the unique solution of the implicit system on $x^0 + (I_1 \times I_2 \times \ldots \times I_{d-l})$.

Arbitrary Dimension

Notice that the last proposition is valid just for C^1 assumptions. We also get an evaluation of the existence neighborhood.

In particular we also get that the differential system has unique solution although the right-hand side is just continuous.

Again by the implicit functions theorem and by the relation

$y_{d-l}(t_1, t_2, \ldots, t_j, 0, \ldots, 0) = y_j(t_1, t_2, \ldots, t_j)$

we see that $y_1, y_2, \ldots, y_{d-l}$ have continuous partial derivatives with respect to their arguments.

IMPORTANT: the above parametrizations may be more advantageous in applications since we may use maximal solutions, as it was shown in dimension three.

Generalized Solutions in Arbitrary Dimension

We discuss now the nonlinear implicit system in the absence of the nondegeneracy hypothesis, i. e. for $\det A(x^0) = 0$ (in fact all the maximal order determinants are null). We consider instead that there is $\{x^n\} \subset \Omega$, such that:

$$x^n \to x^0, \quad \mathrm{rank} J(x^n) = l,\ n \in N,$$

where $J(x^n)$ denotes the Jacobian matrix of $F_1, F_2, \dots, F_l \in C^1(\Omega)$, in x^n.

Notice that in case this is not fulfilled, it means that rank $J(x) < l$ in $x \in W$, where W is a neighbourhood of x^0. Then $F_1, F_2, \dots, F_l$ are not functionally independent in W and the problem (1.1) can be reformulated by using less functionals. That is the above property is in fact always valid, except for not well formulated implicit systems.

Generalized Solutions in Arbitrary Dimension

In each x^n, one can solve the system $F_j(x) = F_j(x^n),\ j = \overline{1,l},\ x \in \overline{\Omega}_1$, where Ω_1 is open and bounded such that $x^0 \in \Omega_1 \subset\subset \Omega$ and one can find the solution via the differential system.
We denote by $T_n \subset R^d$ the above solution and it may be assumed a compact in R^d (it is clearly closed due to the continuity of $F_j, j = \overline{1,l}$). We may also assume that $\{T_n\}$ are uniformly bounded since Ω_1 is bounded and, on a subsequence α, we have $T_n \to T_\alpha,\ n \to \infty$, in the Hausdorff-Pompeiu metric, where T_α is some compact subset in R^d.

Definition

$T = \bigcup_\alpha T_\alpha$ is the local generalized solution of the nonlinear system in x^0, in the critical case rank $J(x^0) < l$. The union is taken for all the sequences and subsequences satisfying the above conditions.

Generalized Solutions in Arbitrary Dimension

The above definitions cover all the possible critical or non critical cases. For instance, if we have just one equation and x^0 is an extremum for the respective function, then the generalized solution is just x^0. If the respective function is identically zero in $O \subset \Omega$ and x^0 is on the boundary of O, then the generalized solution is the boundary of O - see the Example below. A complete description of the level sets (even of positive measure) around x^0 may be obtained via the generalized solution. Generally speaking, the generalized solution is not a manifold and may be not a compact subset. The computed approximation may be not connected.

Generalized Solutions in Arbitrary Dimension

Proposition

We have $x^0 \in T_\alpha \subset T \subset \partial M_{x_0}, \forall \alpha$, where ∂M_{x_0} is the connected component of ∂M containing x_0 and M is the solution. In particular

$$F_j(x) = 0,\ j = \overline{1, l},\ \forall\, x \in T.$$

If x^0 is a regular point, then we denote by S the (local) solution obtained via the implicit function theorem around x^0. In the *Definition*, we choose $x^n \to x^0$, $x^n \in S$ and the uniqueness property from the implicit functions theorem gives that $T_n = S$, locally for n big enough. This choice of x^n satisfies the conditions since $J(x^n) \to J(x^0)$.

We see that in the classical case, one obtains $T = S$ (locally), that is the *Definition* gives indeed a generalization of the classical local solution of the implicit functions theorem.

Generalized Solutions in Arbitrary Dimension

In R^2, take $d = 2, l = 1$ and

$$f(x_1, x_2) = \begin{cases} x_1^2(x_2^2 - x_1^2)^2 & \textit{if } x_1 < 0, \quad |x_2| \le |x_1| \\ 0 & \textit{otherwise.} \end{cases}$$

Clearly f is in $C^1(R^2)$ and $\nabla f(x_1, x_2) = 0$, on the second line.
Take $x^0 = (0, 0)$ and $x^n \to x^0$, $x^n = (x_1^n, x_2^n)$, $x_1^n < 0$, $|x_2^n| < |x_1^n|$.
In such points x^n, one can use previous theorems and the differential system may be chosen of Hamiltonian type:

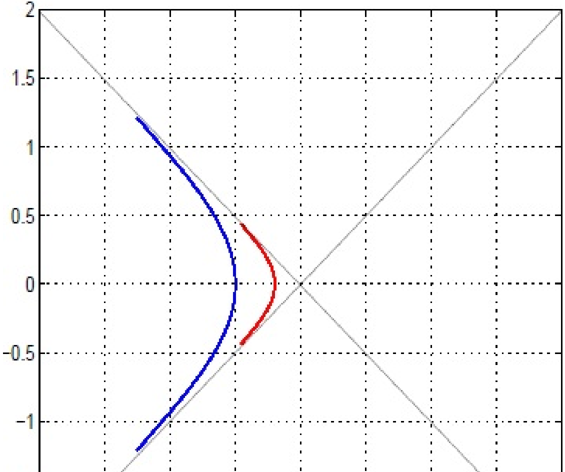
$$x_1'(t) = -4x_1^2 x_2(x_2^2 - x_1^2),$$

$$x_2'(t) = 2x_1(x_2^2 - x_1^2)(x_2^2 - 3x_1^2),$$

$$(x_1(0), x_2(0)) = x^n.$$

Generalized Solutions in Arbitrary Dimension

We represent the solution T_n obtained with Matlab, for $x^n = (-\frac{1}{n}, 0), n = 2, 5$



Generalized Solutions in Arbitrary Dimension

This generalized solution contains the essential information about the solution set of (1.1), since it gives its boundary (and in the proposition, the inclusion becomes equality). If we define

$$f_1(x_1, x_2) = x_1^2(x_1^2 + x_2^2 - 1)_-^2$$

and $x^0 = (0, 1)$, then ∂M is connected and the corresponding generalized solution is ∂M without the lower half of the unit circle. The inclusion is strict in this case. This is also related to the local character of our construction.

Proposition

Let x^0 be the unique critical point of (1.1) in the closed ball $B(x^0)$. Then, $T = M$ in $B(x^0)$.

Generalized Solutions in Arbitrary Dimension

Proposition

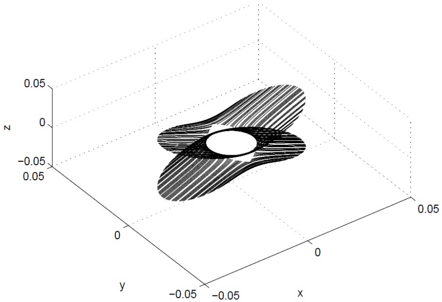
Let $F_j \in C^1(\Omega)$, $j = \overline{1,l}$ *and* $x^n \to x^0$, $x^n, x^0 \in \Omega$. *Denote by* $\widetilde{T}_n, \widetilde{T}_0$ *the generalized solutions of (1.1) contained in the bounded domain* Ω, *corresponding to the initial conditions* x^n, *respectively* x^0. *Then*

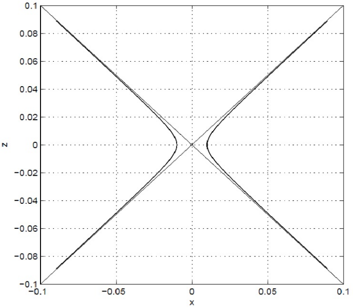
$$\limsup_{n\to\infty} \widetilde{T}_n \subset \widetilde{T}_0.$$

There are examples with the above inclusion strict.

DIMENSION THREE : EXAMPLES CRITICAL CASE

$f(x, y, z) = x^2 + y^2 - z^2$, $(x_0, y_0, z_0) = (0, 0, 0)$,
$\nabla f(x_0, y_0, z_0) = (0, 0, 0)$





The generalized solution around (x_0, y_0, z_0) and its section through a vertical plane. Here, we have $T = M$.

Applications

One motivating application for the above discussion is in shape optimization problems. A typical example has the form:

$$\operatorname*{Min}_{\Omega \in \mathcal{O}} \int_{\Lambda} j(x, y_\Omega(x))dx,$$

$$-\Delta y_\Omega = f \quad in\ \Omega,$$

$$y_\Omega = 0 \quad on\ \partial\Omega$$

with other supplementary constraints, if necessary. Other differential operators or other boundary conditions may be studied as well.

Applications

Here, $\mathcal{O}$ is a family of admissible domains in R^3, satisfying certain regularity hypotheses and conditions like

$$E \subset \Omega \subset D, \ \forall\, \Omega \in \mathcal{O}$$

where E, D are given (bounded) domains, E may be even void, etc.

Function $f \in L^2(D)$ and Λ may be either E (if nonvoid) or Ω or $\partial\Omega$, etc. The integrand $j(\cdot, \cdot) : D \times R \to R$ is of Carathéodory type.

The problem has a similar structure with an optimal control problem, the main difference and difficulty being that the minimization parameter is the domain Ω itself.

Applications

The idea is to represent the unknown domains as level sets $F(x_1, x_2, x_3) \leq 0$ in D. We assume as usual $F_{,1}(x^0) \neq 0$.

Now, to compute the gradient of such a cost functional, we can use functional variations that are perturbations of the form

$$F(x_1, x_2, x_3) + \lambda h(x_1, x_2, x_3) = 0,$$

Such perturbations of the geometry may be very complex, including topological and boundary perturbations simultaneously. This is important in applications to shape optimization .

Applications

We approximate the equation in D

$$-\Delta y_\varepsilon + \frac{1}{\varepsilon}(1 - H^\varepsilon(g))y_\varepsilon = f \quad \text{in } D,$$

$$y_\varepsilon = 0 \quad \text{on } \partial D.$$

Proposition *If $\Omega = \Omega_g$ is of class C, then $y_\varepsilon|_{\Omega_g} \to y$ (the solution of the original state system) weakly in $H^1(\Omega_g)$ and strongly in $L^2(\Omega_g)$.*

Applications

An example in dimension two is the following

$$\operatorname*{Min}_{\Omega\in\mathcal{O}} \int_{\Lambda} \left(\frac{\partial y_\Omega}{\partial n}\right)^2 d\sigma,$$

$$-\Delta y_\Omega = f \quad \text{in } \Omega,$$

$$y_\Omega = 0 \quad \text{on } \partial\Omega,$$

$$E \subset \Omega \subset D, \ \forall\, \Omega \in \mathcal{O}.$$

We assume that the admissible domains in $\mathcal{O}$ are defined as level sets of functions $g \in G_{ad}$.

Applications

Proposition *If* $f \in L^p(D), p > 2$*, the directional derivative of the cost is given by*

$L = \int_I [A + B + C]$,

where

$$A = 2\frac{\partial y_\varepsilon}{\partial n_g}(x_g(t), y_g(t))\{\frac{\partial}{\partial n_g}[\nabla y_\varepsilon(x_g(t), y_g(t))].(z(t), w(t)) + [\nabla y_\varepsilon(x_g(t), y_g(t)$$

$$[\nabla n_g(x_g(t), y_g(t)).(z(t), w(t))]\}\sqrt{(\dot{x}_g(t))^2 + (\dot{y}_g(t))^2},$$

Applications

$$B = 2\frac{\partial y_\varepsilon}{\partial n_g}(x_g(t), y_g(t))\{[\nabla y_\varepsilon.\nabla h/|\nabla g|](x_g(t), y_g(t)) - [\frac{\partial y_\varepsilon}{\partial n_g}\nabla g.\nabla h/|\nabla g|^2]$$

$$(x_g(t), y_g(t))\}\sqrt{(\dot{x}_g(t))^2 + (\dot{y}_g(t))^2},$$

$$C = 2\{[\frac{\partial y_\varepsilon}{\partial n_g}^2 |\nabla g|^{-2}](x_g(t), y_g(t))\}(\dot{x}_g(t), \dot{y}_g(t)).(z(t), w(t)).$$

where (z, w) solve the system in variations associated with the Hamiltonian system describing the geometry and I is its existence interval.

Applications

Due to such applications (and again in arbitrary dimension), one can also consider general perturbations having the form

$$F_k^\lambda(x_1, \dots, x_d) = 0, \quad k = \overline{1, l},\ \lambda \in (-1, 1),$$

where $F_k^\lambda \in C^2(\Omega \times (-1, 1))$, $F_k^0 = F_k$ and $F_k^\lambda(x^0) = 0$, $k = \overline{1, l}$, and the independence hypothesis remains clearly valid for the perturbation as well, for λ small.

If the second approach (with implicit functions) is used, we get the following result, where $z_j^\lambda(t_1, \dots, t_j)$ denote the corresponding variations of the solutions of the implicit system. The existence of $z_j^\lambda(t_1, \dots, t_j)$ follows as before.

Applications

Proposition

Assume that $F_k^\lambda \in C^1(\Omega \times (-1,1))$. We have:
a) the last $d-l$ components of $z_1^\lambda(t_1), \dots, z_{d-l}^\lambda(t_1, t_2, \dots, t_{d-l})$ are null.
b) for any $j = 1, \dots, d-l$ and $(t_1, t_2, \dots, t_{d-l}) \in I_1 \times \dots \times I_{d-l}$, $z_j^\lambda(t_1, t_2, \dots, t_j)$ is the unique solution of:

$$\nabla_y F_k^\lambda(y_j^\lambda) z_j^\lambda + \partial_\lambda F_k^\lambda(y_j^\lambda) = 0, \; k = 1, \dots, l.$$

One can obtain for $z_1^\lambda(t_1), \dots, z_{d-l}^\lambda(t_1, t_2, \dots, t_{d-l})$ similar relations even for implicit parametrizations, but point *a)* is not valid and the algebraic system is not uniquely determining them. To overcome this difficulty, one has to use directly the differential systems.

Applications

We consider now applications to the classical nonlinear programming problem, first just with equality constraints:

$$(P) \qquad Min\{g(x_1, \ldots, x_d)\}$$

subject to the general implicit system. Notice that our independence hypothesis is exactly the form of the classical Mangasarian-Fromowitz condition for the special case of problem (P). By Theorem on implicit systems we can replace it by the unconstrained problem for $(t_1, t_2, \ldots, t_{d-l}) \in (I_1 \times I_2 \times \ldots \times I_{d-l})$:

$$(P_1) \quad Min\{g(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)\}$$

Applications

where $(y^1_{d-l}, y^2_{d-l}, \dots, y^l_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \dots, t_{d-l} + x^0_d)$ are the components of y_{d-l}, the solution of the differential system corresponding to this case. This is done around the point x^0, which is mapped into $(0, \dots, 0)$ via the definition of (P_1).

Proposition

Let g and $F_i, i = \overline{1, l}$, be in $C^1(R^d)$ and independence condition be valid in some point x^0. Then, the gradient in $(0, \dots, 0)$ of the composed cost functional in (P_1) has the components:

$$\sum_{i=1}^{l} \frac{\partial g(x^0)}{\partial i}\; v^i_j(x^0) + \frac{\partial g(x^0)}{\partial (j+l)}, \quad j = \overline{1, d-l},$$

Applications

This methodology can be extended to the case of implicit parametrizations, important in the sequel. The form of the above partial derivatives is due to the choice of the last $d - l$ components of the solutions and we can write shortly

$$\frac{\partial}{\partial t_j} G(0, 0, \dots, 0) = \nabla g(x^0).v_j(x^0) \;\; j = \overline{1, d - l}. \tag{1}$$

Notice that it is sufficient to solve just the $l \times d$ linear algebraic system. We obtain the first order optimality conditions in the Fermat form:

Corollary

If x^0 is a local solution of (P) satisfying that g and $F_i, i = \overline{1, l}$, are in $C^1(R^d)$ and independence holds, then we have:

$$\nabla g(x^0).v_j(x^0) = 0 \;\; j = \overline{1, d - l}. \tag{2}$$

Applications

In fact this is equivalent with the classical Lagrange multipliers rule, since it produces a basis in the tangent space at x^0 to the manifold defined by the system. Then this shows that $\nabla g(x^0)$ is in the normal space, which has the basis given by $\nabla F_i(x^0), i = \overline{1, l}$, due to hypothesis. Consequently, there are scalars λ_i such that the well known relation $\nabla g(x^0) = \sum_{i=1}^{l} \lambda_i \nabla F_i(x^0)$ holds. The argument works in the converse sense too. We also underline the explicit character of our result in comparison with the Lagrange rule and its applicability to numerical minimization routines of gradient type. The vector appearing in the right-hand side is called the tangential gradient of g in x^0, to the manifold defined by the system under independence condition.

Applications

We discuss now the general case of both equality and inequality constraints:

$(Q) \qquad Min\{g(x_1, \dots, x_d)\}$

subject to the system and to

$$G_j(x) \le 0 \;\; j = \overline{1, m}, \tag{3}$$

where g, F_i, G_j are in $C^1(R^d)$. The Mangasarian-Fromowitz condition in this case consists of independence and there is $d \in R^d$ such that

$$\nabla F_i(x^0)d = 0, \;\; i = \overline{1, l}, \;\; \nabla G_j(x^0)d < 0, \;\; j \in I(x^0), \tag{4}$$

with $I(x^0)$ being the set of indices of active inequality constraints in x^0.

Applications

The reduced problem is again obtained via parametrization Theorem :

(Q_1) $Min\{g(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)\}$,

subject to the constraints, in the "reduced" form:

$$G_j(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1+x^0_{l+1}, \ldots, t_{d-l}+x^0_d) \leq 0 \quad j = \overline{1, m}, \quad (5)$$

Applications

Lemma

The minimization problem (Q_1) *satisfies the Mangasarian-Fromowitz condition in the origin of* R^{d-l}.

If x^0 is a local solution of (Q), by Lemma , one can apply the classical KKT theorem to the problem (Q_1) in the origin of R^{d-l} that is a local solution. Using again the derivation formula , we get:

Theorem

Let x^0 *be a local minimum for* (Q). *Then, there are* $\beta_j \geq 0, j = \overline{1,m}$ *such that*

$$0 = \nabla g(x^0).v_s(x^0) + \sum_{j=1}^{m} \beta_j \nabla G_j(x^0), s = \overline{1, d-l},$$

$$0 = \beta_j G_j(x^0), j = \overline{1,m}.$$

Applications

This is a simplified version of the KKT conditions since it eliminates the Lagrange multipliers for the equality constraints. Notice that the statement is in terms of data, i.e. the problem (Q) and the differential system is not explicitly used in previous Theorem. The reduced problem is $d - l$ dimensional. Taking into account that just the multipliers associated with the active inequality constraints are non zero, then we think to eliminate completely the multipliers in this case too.

We relax now the hypotheses in the problem (Q) and we describe a direct minimization algorithm.

Applications

It looks for the solution in a maximal neighborhood of x^0.
We assume in the sequel that g and $G_j, j = \overline{1, m}$, are just in $C(R^d)$ and $F_i, i = \overline{1, l}$, are in $C^1(R^d)$ and satisfy independence condition in x^0. This last condition can be removed in fact, via generalized solutions.
Notice that x^0 is here just an admissible point for (Q) and not a local minimum. We can also add the abstract constraint $x \in D$, some given subset in R^d, such that $x^0 \in D$.

Applications

ALGORITHM

1)choose $n = 1$, the discretization step $1/n$ and the solution intervals $I_1^n, \dots, I_{d-l}^n$, the tolerance δ, etc.
2)compute the discrete set of admissible points C_n, starting from x^0, via previous systems
3)find in C_n the approximating minimum of (Q), denoted by x^n
4) test if the solution is satisfactory by $|g(x_n) - g(x_{n-1}| \leq \delta$.
5) If YES, then STOP. If NO, then $n := n + 1$ and GO TO step 1).

Applications

In step 4) other tests (on the solutions, on the gradients, etc.) may be used. The approximating solution $x^n \in C_n$ may be not unique and one can adapt the convergence test to such situations.

Theorem

The algorithm is convergent as $n \to \infty$.

Notice the global character of this approach. In case there are several global minimum points in the searched region, then the algorithm will find all of them.

Applications

The set defined by the equality constraints may have several connected components. Starting from x^0, Algorithm will minimize just on the component that contains x^0. Initial guesses from all the admissible components are necessary if we want to minimize on all of them.
If independence condition is not fulfilled , then one can use the generalized solution of implicit systems as explained in Section 3, since the Hausdorff-Pompeiu distance ensures the uniform convergence of approximating points. The obtained minimum may satisfy constraints and the minimum property with some small error tolerance.

Applications

Finally, we comment some illustrative numerical examples. We consider first a minimization problem on the torus in R^3, with radiuses 2 and 1 and with initial point $(x_0, y_0, z_0) = (\sqrt{5}, 2, 0)$:

$$min\{xyz\}$$
$$F(x, y, z) = (x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2)$$

The obtained results are given below, compared with the application of the fmincon routine of MatLab:

$$min = -2,7154$$
$$x_{min} = 1,7841; y_{min} = 1,8199; z_{min} = -0,8363$$
$$fmincon : min = -2,7153; x_{min} = 1,802; y_{min} = 1,802; z_{min} = -0.$$

Using other starting points like $(1, 0, 0)$ or $(3, 0, 0)$ is not allowed by MatLab that finds no admissible solutions in these cases, while our approach works.

Applications

Now, we consider two equality restrictions, given by F and P. Two initial points are taken into account since the intersection has two components.

$$min\{x^3 + 5y - 7sinz\}$$

$$P(x,y,z) = \frac{2\sqrt{3}}{3}x - y^2 - z^2$$

$$(x_0, y_0, z_0) = (\sqrt{3}, 1, 1);\ (x_0, y_0, z_0) = (\sqrt{3}, -1, 1)$$

The numerical results and a comparison with MatLab:

$$(\sqrt{3}, -1, 1) : minimal\ value = 0.498975897823261$$

$$solution : \ (1.06688905550184, -0.814925789648031, 0.7536319$$

$$(\sqrt{3}, 1, 1) : minimal\ value = -7.65929313197537$$

$$solution : \ (1.10697710321061, 0.817093948780941, 0.7814791 2$$

Applications

In the second case fmincon stops after 42 iterations with the message that constraints are not satisfied within the tolerance. In the first case, fmincon finds basically the same solution.

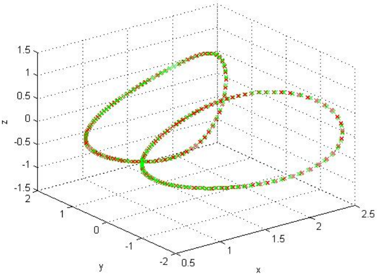


Figure: The admissible set

Applications

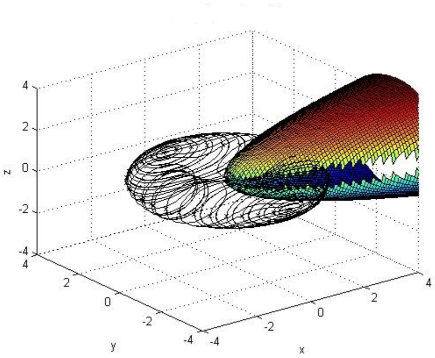


Figure: The geometry

Applications

In M.D. Stuber, J.K. Scott, P.I. Barton, Convex and concave relaxations of implicit functions, Optimization methods and software, 30(3), pp.424-460 (2015), an example in R^6, with three equality constraints, is discussed. Reworking it via our Algorithm , starting from the two solutions indicated there on p.451, we obtain the new points (0,5631;-3,2581; 0:51593; 0,4692; 1,4635; 3,589), (0,56166;-3,3154; 0,50897; 0,5047; 1,4365; 3,6777)
with the cost values 343,7695, respectively 383,7265. This improves the quoted experiment and can be directly checked. But there is no contradiction since we increase the search region.

THANK YOU!