On a Geometrical Interpolation Problem

Octavian Stănăşilă

Abstract

Given a finite set of points A in $\mathbb{R}^n$ and a geometrical pattern Φ we define a type of distance between A and Φ and study how to find a pattern Φ among a parametrized family of geometrical objects such that such a distance is minimum.

Mathematics Subject Classification: 41A29
Key words: pattern, distance, point–pattern, least squares interpolation problem, Newton–Raphson method.

1 Preliminaries

Many bidimensional "patterns" have a concatenation of plane curve arcs as boundaries; similarly, 3-dimensional "patterns" have as boundary a concatenation of surface pieces. For this reason, it could be useful to introduce and to study some new metrical properties, which permit to formulate and solve some interpolation problems.

Let $U \subset \mathbb{R}^n$ be an open set and fix an integer $k \geq 1$. Consider a map $f : U \to \mathbb{R}^k$ of class $C^1(U)$, $f = (f_1, ..., f_k)^T$; for any $x \in U$, denote $J_f(x) = \left(\frac{\partial f_i}{\partial x_j}\right)$; $1 \leq i \leq k$, $1 \leq j \leq n$, the Jacobian matrix. The most important case in what follows will be when $k < n$; in this case, a point $x \in U$ is said regular for f if $rank(J_f(x)) = k$. The set of type $f^{-1}(0)$ generalize the plane curves, the surfaces, space curves etc.

Definition 1.1. We call a *pattern* in U any set of the form $\Phi = \bigcup_{i=1}^{M} \Phi_i$, such that there are C^1 functions $g_i : U \to \mathbb{R}^k$, $1 \leq i \leq M$ and $\Phi_i \subset g_i^{-1}(0)$ have all their points regular.

If $k = 1$, $\Phi \subset g^{-1}(0)$, where $g = g_1 \cdot g_2 ... g_M$; if x is regular for g_i, $1 \leq i \leq M$, it could be not regular for g.

Given a point $a \in U$, $a = (a_1, ..., a_n)^T$ a pattern Φ as above, it could be useful to define a suitable distance $d(a, \Phi)$ between them. If so, take $A_1, A_2, ..., A_N \in U$ as N distinct points ("points of surveillance"); one knows that even the simple interpolation Lagrange problem has some obstructions to be solved, instead it

Balkan Journal of Geometry and Its Applications, Vol.1, No.2, 1996, pp. 97-103

could be useful to determine, under some conditions, a function f such that the sum $\sum_{i=1}^{N} d(A_i,\ f^{-1}(0))^2$ is minimum. Such problems will be treated in what follows.

On the other hand, recall that if $A \in M_{k,n}(\mathbb{R}), B \in M_{k,1}(\mathbb{R})$ are two matrices, then $A^+ \in M_{n,k}(\mathbb{R})$ means the Penrose pseudoinverse of A, [4] and by putting $B^+ = AA^+B$, the vector $\xi = A^+B$ is unique in $\mathbb{R}^n$ such that $A\xi = B^+$, the latter being just the orthogonal projection of B on ImA (that is $\| B - B^+ \| =$ minimum); ξ is called the *pseudosolution* of the linear system $AX = B$. If $k \le n$ and $rankA = n$, then $A^+ = A^T \cdot (AA^T)^{-1}$.

2 A distance between a point and a pattern

Definition 2.1. Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^k$, $f = (f_1, ..., f_k)^T$ be a C^1-map; for any point $a \in U$ define the *distance* from a to the pattern $f^{-1}(0)$ as being

$$\delta(a, f^{-1}(0)) = \| J_f(a)^+ \cdot f(a) \| \text{ (Euclidean norm)} \tag{1}$$

Example. In the case when $k = 1$ and f is linear nonnull, $f(x) = \sum_{i=1}^{n} c_i x_i$, put $c = (c_1, ..., c_n)^T$; then for any $a \in \mathbb{R}^n$, $a = (a_1, ..., a_n)^T$, one has $f(a) = c^T \cdot a$, $J_f(a) = c^T$ and $J_f(a)^+ = \frac{c}{\|c\|^2}$. Put $p = a - J_f(a)^+ \cdot f(a)$ hence $p = a - \frac{1}{\|c\|^2}(c \cdot c^T \cdot a)$. Then $f(p) = c^T \cdot p = 0$ and the vector $a - p$ is normal to the hyperplane $f^{-1}(0)$. In this case, $\| a - p \| = \delta(a, f^{-1}(0))$, justifying thus the definition 2.1. So δ extends the Euclidean distance.

Proposition 2.2. *Suppose that $1 \le k < n$ and $a \in U$ is a regular point for a C^1-map $f : U \to \mathbb{R}^k$. Then*

$$\delta(a, f^{-1}(0)) = \left(f(a)^T \cdot \left(J_f(a) \cdot J_f(a)^T \right)^{-1} \cdot f(a) \right)^{\frac{1}{2}}. \tag{2}$$

Proof. Let $J = J_f(a)$ hence rank$J = k$ (maximum); in this case, the symmetrical matrix $J \cdot J^T$ is invertible and moreover, $J^+ = J^T \cdot (J \cdot J^T)^{-1}$. Then by the definition 2.1.,

$$\delta(a, f^{-1}(0))^2 = \| J^+ \cdot f(a) \|^2 = < J^+ \cdot f(a), J^+ \cdot f(a) > = f(a)^T \cdot (J^+)^T \cdot J^+ \cdot f(a).$$

Since $J^+ = J^T(J \cdot J^T)^{-1}$, one obtains:

$$\delta(a, f^{-1}(0))^2 = f(a)^T \cdot (J \cdot J^T)^{-1} \cdot J \cdot J^T \cdot (J \cdot J^T)^{-1} \cdot f(a).$$

But $J \cdot J^T \cdot (J \cdot J^T)^{-1} = I_n$ therefore $\delta(a, f^{-1}(0))^2 = f(a)^T \cdot (J \cdot J^T)^{-1} \cdot f(a)$, whence the proposition.

Corollary 2.3. *Fix $a \in U$. The map $f \to \delta(a, f^{-1}(0))$, restricted to C^1–functions on U for which a is regular, is continuous.*

Corollary 2.4. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $f_i(x) = \sum_{j=1}^{n} a_{ij} x_j$, where $\sum_{j=1}^{n} a_{ij}^2 = 1$ for $1 \le i \le k < n$ and the $k \times n$-matrix $A = (a_{ij})$ has the rank k. Consider the hyperplanes*

$$H_i = \{x \in \mathbb{R}^n | f_i(x) = b_i\}, \quad 1 \le i \le k$$

and put $B = (b_1, ..., b_k)^T$*. Then the pseudosolution of the linear system* $A \cdot X = B$ *minimizes the sum* $\sum_{i=1}^{k} \delta(x, H_i)^2$.
Proof. Define $f : \mathbb{R}^n \to \mathbb{R}^k$, $f(x) = (f_1(x) - b_1, ..., f_k(x) - b_k)$ hence by proposition 2.2 $\delta(x, f^{-1}(0)) = f(x)^T \cdot (AA^T)^{-1} \cdot f(x)$ for any $x \in \mathbb{R}^n$. Put $B^+ = AA^+B = (b_1^+, ..., b_k^+)^T$; then $\xi = A^+B$ is the pseudosolution of the system $AX = B$ and $A\xi = AA^+B = B^+$. We have

$$\delta(x, H_i) = d(x, H_i) = |f_i(x) - b_i|$$

and

$$\delta(\xi, H_i) = d(\xi, H_i) = |f_i(\xi) - b_i|,$$

hence

$$\sum_{i=1}^{k} \delta(x, H_i)^2 = d(f(x), b)^2 \geq d(Imf, b)^2 \geq \| B - B^+ \|^2 =$$

$$= \sum_{i=1}^{k} (b_i^+ - b_i)^2 = \sum_{i=1}^{k} |f_i(\xi) - b_i|^2 = \sum_{i=1}^{k} \delta(\xi, H_i)^2.$$

The distance given in the definition 2.1 has some convenient geometrical properties. Under obvious hypothesis, one directly proves:
Proposition 2.5. *Let* $a \in U$ *and a* C^1*-function* $f : U \to \mathbb{R}^k$.
1) *If* $A \in M_k(\mathbb{R})$ *is nonsingular and* $g = A \cdot f$*, then* $\delta(a, f^{-1}(0)) = \delta(a, g^{-1}(0))$;
2) *If* $Q \in M_n(\mathbb{R})$ *is an orthogonal matrix and* $T : \mathbb{R}^n \to \mathbb{R}^n$, $Tx = Qx + c$ *is an isometry, then* $\delta(a, f^{-1}(0)) = \delta(Ta, Tf^{-1}(0))$;
3) *For any* $\lambda > 0$, $\delta(\lambda a,\ g^{-1}(0)) = \lambda\delta(a, f^{-1}(0))$, *where* $g(x) = f\left(\frac{x}{\lambda}\right)$.

3 A least-square interpolation problem and an algorithm

Let $\mathcal{A} = \{A_1, ..., A_N\} \subset U$ $(U \subset \mathbb{R}^n$open) be a *fixed* set of distinct "points of surveillance". For any pattern $\Phi = f^{-1}(0)$, where $f : U \to \mathbb{R}^k$ is a C^1-function such that each A_i is regular for f, we put

$$\delta(\mathcal{A}, \Phi) = \sum_{i=1}^{N} \delta(A_i, f^{-1}(0))^2. \tag{3}$$

One can formulate the following problem:
(Π). *Given* $\mathcal{A}$*, determine a function* f *which minimizes* $\delta(\mathcal{A}, f^{-1}(0))$. Such a problem could have some applications in Pattern Recognition [3]. In fact it is a geometrical nonlinear variant of the least squares method and the solution is generally not unique. In [1] one proves that whenever $F \subset \mathbb{R}^n$ is closed and $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, with $F = f^{-1}(0)$, then for any continuous function $\epsilon : \mathbb{R}^n \to (0, \infty)$ there exists a C^∞-function $g : \mathbb{R}^n \to \mathbb{R}$ such that $g^{-1}(0) = F$ and $\forall x \in \mathbb{R}^n$, $|f(x) - g(x)| < \epsilon(x)$. This shows that for any pattern there are smooth functions which define it.

In what follows, we present the case when the unknown belongs to a parametrized family of functions (e.g. quadrics, cubics, spline-functions etc).

Let $P \subset \mathbb{R}^m$ be an open subset in a parameters space and $F : U \times P \to \mathbb{R}^k$ be a map such that for any $p \in P$, $p = (p_1, ..., p_m)^T$, F determines a C^1-function $f : U \to \mathbb{R}^k$, $x \to F(x, p)$; suppose there is a bijective correspondence between such functions and parameters (this happens for instance in the case of the polynomial functions of degree at most d, where $P = \mathbb{R}^n$ and $m = \begin{pmatrix} n+d \\ n \end{pmatrix}$; in such a case, the function $\delta(\mathcal{A}, \Phi)$ becomes a function of the coefficients of the polynomials). In the case of the linear dependence on parameters, we also can impose supplementary relations on parameters, which do not modify the solution; for instance, the following condition

$$\sum_{i=1}^{N} J_f(A_i) \cdot J_f(A_i)^T = I_k. \tag{4}$$

Indeed, since the matrices $J_f(A_i) \cdot J_f(A_i)^T$ are symmetrical, positively semi-definite and nonsingular (A_i being supposed regular points of f), these matrixes result positively definite and the same is true for the matrix $C = \sum_{i=1}^{N} J_f(A_i) \cdot J_f(A_i)^T$. So there is an orthogonal matrix $Q \in M_k(\mathbb{R})$ such that $Q^T \cdot C \cdot Q = I_k$ and therefore Qf will verify (4).

Suppose now that the points $A_1, ..., A_N$ are "sufficiently near" of $f^{-1}(0)$, in the sense that $\delta(A_i, f^{-1}(0)) \simeq \| f(A_i) \|$, for each i. By (3), $\delta(\mathcal{A}, f^{-1}(0)) \simeq \sum_{i=1}^{N} \| f(A_i) \|^2$. Finally suppose fixed r C^2-functions $\varphi_1, ..., \varphi_r : U \to \mathbb{R}^k$, linearly independent and consider that the unknown C^2-function f has the form $f = (f_1, ..., f_k)^T$, where $f_i = \sum_{j=1}^{r} p_{ij}\varphi_j, \quad 1 \le i \le k$, with p_{ij} unknown real constant (as parameters); in matricial writing, $f = P^T \cdot \varphi$, where $P = (p_{ij}) \in M_{r,k}(\mathbb{R})$. Check the function f, that is the matrix P, such that the sum $\sum_{i=1}^{N} \| f(A_i) \|^2$ is minimum, with the restriction (4). In this case,

$$\sum_{i=1}^{N} \| f(A_i) \|^2 = \sum_{i=1}^{N} f(A_i)^T \cdot f(A_i) =$$

$$= \sum_{i=1}^{N} \varphi(A_i)^T \cdot P \cdot P^T \cdot \varphi(A_i) = tr(P^T \cdot A \cdot P),$$

where $A = \sum_{i=1}^{N} \varphi(A_i) \cdot \varphi(A_i)^T \in M_r(\mathbb{R})$ is a known symmetrical, positively semi-definite matrix. On the other part, since

$$J_f(A_i) \cdot J_f(A_i)^T = P^T \cdot J_\varphi(A_i) \cdot J_\varphi(A_i)^T \cdot P,$$

by putting

$$B = \sum_{i=1}^{N} J_\varphi(A_i) \cdot J_\varphi(A_i)^T,$$

with $B \in M_r(\mathbb{R})$ known, the relation (4) becomes $P^T \cdot B \cdot P = I_k$. Thus, the above least-squares problem reduces to the following matricial one: *check a matrix* $P \in M_{r,k}(R)$ *of rank* k *such that* $P^T \cdot B \cdot P = I_k$ *and* $tr(P^T \cdot A \cdot P)$ *be minimum.*

In order to solve this, we first remark that $P^T \cdot A \cdot P$ can be assumed diagonal [indeed, $P^T \cdot A \cdot P$ is symmetrical and positively semi-definite, hence there is an orthogonal matrix $Q \in M_k(\mathbb{R})$, i.e., $Q \cdot Q^T = I_k$, such that $Q^T(P^T \cdot A \cdot P)Q$ is diagonal; but $tr(P^T \cdot A \cdot P) = tr((PQ)^T \cdot A \cdot P \cdot Q)$ and moreover, $P^T \cdot B \cdot P = I_k$ if and only if $(PQ)^T \cdot B \cdot P \cdot Q = I_k$. Thus, one can substitute P by PQ]. Denote by $x_1, ..., x_k$ the column vectors of the matrix P; then $tr(P^T \cdot A \cdot P) = \sum_{i=1}^k x_i^T A x_i$ and the relation $P^T \cdot B \cdot P = I_k$ can be written $x_i^T \cdot B \cdot x_j = \delta_{ij}$ for any $1 \le i, j \le k$. Consider the Lagrangean

$$\mathcal{L}(x_1, ..., x_k, L) = \sum_{i=1}^k x_i^T A x_i - \sum_{u=1}^k \sum_{v=1}^k \lambda_{uv}(x_u^T B x_v - \delta_{uv}),$$

where $L = (\lambda_{uv})$; $1 \le u, v \le k$ is the matrix of the Lagrange multipliers; the necessary (here also sufficient) extreme condition is $\nabla \mathcal{L} = 0$ and this leeds to $AP - BPL = 0$. But $P^T AP = D$ (diagonal) hence $L = I_k L = (P^T BP)L = P^T(BPL) = P^T AP = D$. Therefore the matrix L will be diagonal and much more, positively semi-definite (since A is so). Put $L = diag(\alpha_1, ..., \alpha_k)$ and the relation $AP - BPL = 0$ leeds to $(A - \alpha_i B)x_i = 0$, $1 \le i \le k$, hence x_i is an eigenvector of the matrix-bundle $A - \alpha B$, with an eigenvalue α_i. One can apply different methods to determine the vectors x_i and, by this, the matrix P.

All the above can be shortly concentrated in the following.

Proposition 3.1. (Algorithm to solve the problem Π). *Let* $U \subset \mathbb{R}^n$ *be an open set. Fix* r *linearly independent* C^2*-functions* $\varphi_1, ..., \varphi_r : U \to \mathbb{R}^k$ *and* N *"points of surveillance"* $A_1, ..., A_N \in U$. *The problem is to look for a function* $f : U \to \mathbb{R}^k$ *[whose components* $f_1, ..., f_k$ *are linear combinations of* $\varphi_1, ..., \varphi_r$*, namely* $(f_1, ..., f_k) = (\varphi_1, ..., \varphi_r)P$*, with* $P \in \mathcal{M}_{r,k}(\mathbb{R})$ *a matrix to be determined], such that the pattern* $f^{-1}(0)$ *is the nearest to the set* $\mathcal{A} = \{A_1, .., A_N\}$.

Step I

Determine the $r \times r$*-matrices* $A = \sum_{i=1}^N \varphi(A_i) \cdot \varphi(A_i)^T$ *and* $B = \sum_{i=1}^N J_\varphi(A_i) \cdot J_\varphi(A_i)^T$*, where* $\varphi = (\varphi_1, ..., \varphi_r)^T$.

Step II

The column-vectors $x_1, ..., x_k$ *of the looked for matrix* P *(of rank* k*) are just the eigen vectors of the matrix-bundle* $A - \alpha B$ *and moreover verify* $P^T BP = I_k$. *One determines thus* $x_1, ..., x_k$ *and* P.

As well as the matrices A, B are known (that depending on the choise of the functions $\varphi_1, ..., \varphi_r$), this algorithm requires $O(r^3)$ operations.

4 A Newton-Raphson type result

In the paper [2] it was proved the following result:

Let $D \subset \mathbb{R}^n$ be a nonempty convex bounded set, $f : D \to D$ a C^2-map such that any point of $f^{-1}(0)$ is nonsingular for f. For $u \in D$ fixed, define the function $h_u : D\times[0,1]\to\mathbb{R}^n$, $h_u(x,t) = f(x)+(t-1)f(u)$. Then for a.e. $u\in D$, the set

$$\{(x,t)\in D\times[0,1] | 0\le t\le 1, \quad h_u(x,t)=0\}$$

either consists of a finite number of closed curves in $D\times[0,1]$, or a finite number of arcs in $D\times(0,1)$ with their ends in $D\times\{1\}$ or $D\times\{0\}$, or a finite number of arcs which start from $D\times\{0\}$ and end in $D\times\{1\}$; all these three kinds of curves are disjoint, of class C^1. Therefore, one can found a solution for the equation $f(x)=0$ by following the curve $h_u^{-1}(0)$ which starts from $(u,0)$ for some $u\in D$; such a curve will attain a solution ξ of the equation $f(x)=0$ as soon as $t=1$ is touched.

Let $f: U\to\mathbb{R}^n$ be a C^2-map ($U\subset\mathbb{R}^n$ open) and $h(x,t)$, $h: U\times\mathbb{R}\to\mathbb{R}^n$ a C^2-map such that for any $x\in U$, $h(x,1)=f(x)$ and the equation $h(x,0)=0$ has a solution $u\in U$; for instance, take $h(x,t)=f(x)+(t-1)f(u)$ as above, or $h(x,t)=(1-t)(x-u)+tf(x)$. Suppose that $(\gamma)=h^{-1}(0)$ is a curve in $\mathbb{R}^{n+1}$ which joins $(u,0)$ and a point $(\xi,1)$ such that $f(\xi)=0$. Let $a\in(\gamma)$ be fixed; choose a tangent vector τ in a at (γ) such that $J_h(a)\cdot\tau=0$, $\parallel\tau\parallel=1$ and $det\begin{pmatrix} J_h(a)\\ \tau^T\end{pmatrix}>0$. Then choose a step $p>0$ sufficiently small such that if $b=a+p\tau$, then $h(b)$ is near to 0. Put $c=b-J_h(b)^+\cdot h(b)$. By the definition 2.1, $\delta(b,h^{-1}(0))=\parallel b-c\parallel$; c is just near the point where the hyperplane, passing by b and orthogonal to b, intersects (γ). If u is a point of simple bifurcation for k, then in the neighborhood of u, $h^{-1}(0)$ represents the union of two curves γ_1,γ_2; if we take the arc length s as parameter and $\gamma_1(0)=u$, $\gamma_2(0)=u$, then $det\begin{pmatrix} J_h(s)\\ \dot\gamma(s)^T\end{pmatrix}$ changes its sign in $s=0$ for $\gamma=\gamma_1$ and $\gamma=\gamma_2$ and conversely; both the curves are suitable for the next algorithm.

Proposition 4.1. *Suppose f,h satisfy the conditions of the above formulated result of [2]. Take a solution $u\in D$ of the equation $h(x,0)=0$. Take $a_0=(u,0)$ and apply the described scheme which yields the sequence $(a_k)_{k\ge0}$, where $a_{k+1}=c$ and $a_k=a$ as above. This sequence converges to a point $(\xi,1)$ such that $f(\xi)=0$.*

The proof uses a typical reasoning for the Newton-Raphson method. The new thing is that b can be singular for h and for this reason we use the pseudoinverse; in fact by the Sard theorem almost all points of $\mathbb{R}^n$ are regular values for the map h.

References

[1] F.Broglia, A.Tognoli, *Approximation of C^∞ functions without changing their zero-set*, Annales Inst.Fourier, 39, fasc.3, 611-632, 1989.

[2] S.N.Chow, J.Mallet-Paret, J.A.Yorke, *Finding zeros of maps by homotopy methods*, Math.Programming (A. Bachem ed), Springer Verlag, 15-56, 1983.

[3] V.Neagoe, O.Stănăşilă, *Teoria recunoaşterii formelor*, Ed.Academiei Romane, 1992.

[4] R.Penrose, *A generalized inverse for matrices*, Proc. Cambridge Phil.Soc., 51, 406-413, 1955.

[5] G.Taubin, *Nonplanar curve and surface estimation in 3-space*, Proc. IEEE Conf.Roboties Automation, 1988.

University Politehnica of Bucharest
Department of Mathematics II
Splaiul Independentei 313
77206 Bucharest, Romania