C. R. Acad. Sci. Paris, t. 325, Série I, p. 665-670, 1997
Statistique/*Statistics*
(Théorie des signaux/*Theory of Signals*)

Estimées localement fortement homogènes

Mila NIKOLOVA

UFR Mathématiques et Informatique, Université René-Descartes,
45, rue des Saints-Pères, 75270 Paris Cedex 06.
E-mail : nikolova@math-info.univ-paris5.fr

Résumé. Cette Note est consacrée à l'estimation régularisée de signaux comportant des zones fortement homogènes – localement constantes, localement linéaires, etc. L'estimée, définie au sens du MAP, est le minimiseur d'une énergie qui est composée d'un terme de fidélité aux données quadratique et d'un terme *a priori* de régularisation. Notre résultat principal est que des zones fortement homogènes sont à la fois retrouvées et conservées localement par l'estimateur, c'est-à-dire qu'elles restent inchangées lorsque les données varient dans un voisinage, si et seulement si la fonction de régularisation est non dérivable sur ces zones.

Locally strongly homogeneous estimates

Abstract. *This Note concerns regularized estimation of signals comprising strongly homogeneous zones* – e.g., *locally constant, locally linear, etc. The estimate, defined in the MAP sense, is the minimizer of an energy which combines a quadratic data-fidelity term and a regularization prior term. Our main result is that strongly homogeneous zones are both recovered and locally conserved by the estimator,* i.e. *that they remain unchanged for data varying in a neighbourhood, if and only if the regularization function is nonsmooth over these zones.*

Abridged English Version

We are concerned with the estimation of an unknown signal $\boldsymbol{x} \in \mathbb{R}^M$, containing large strongly homogeneous zones, from noisy data $\boldsymbol{y} = \widetilde{\mathcal{A}}\boldsymbol{x} + \boldsymbol{n} \in \mathbb{R}^N$, where $\widetilde{\mathcal{A}}$ is a known linear operator and $\boldsymbol{n}$ is white Gaussian noise. The inverse solution $\hat{\boldsymbol{x}}$ is defined as a MAP estimate. It minimizes the posterior energy $\mathcal{E}(\boldsymbol{x})$ in (1-3) involving a *potential function* (PF) φ which is symmetric, increasing on $[0, \infty[$, piecewise $\mathcal{C}^1$, and possibly nonconvex. Implicit functions of the form of $\hat{\boldsymbol{x}}(\boldsymbol{y})$ arise in different settings, such as Bayesian estimation (*see* [2]) and regularization in statistics (*see* [11] and [4]) and in optimization (*see* [12] and [8]).

Our problem is to obtain estimates which are *strongly homogeneous* over large zones, in spite of the noise. Mathematically, a *set of strong homogeneity* is defined as $\mathcal{J} := \{k \ : \ \boldsymbol{d}_k \hat{\boldsymbol{x}} = 0\}$, so it is an implicit function of data, $\mathcal{J} \equiv \mathcal{J}(\boldsymbol{y})$. When φ is nonconvex, $\mathcal{E}_y$ has a set of local minima and each of them gives rise to a set $\mathcal{J}$, as suggested in (5). Statements in this Note concern the strict

Note présentée par Yves MEYER.

local minima of $\mathcal{E}_y$. Henceforth, $\hat{x}$ and $\mathcal{J}$ denote any local minimizer of $\mathcal{E}_y$ and its set of strong homogeneity, respectively.

Let minimizer $\hat{x}$ involve a set of strong homogeneity $\mathcal{J}$; the relevant estimator $\hat{x}(y)$ is said to be *locally strongly homogeneous* if the same $\mathcal{J}$ can be obtained from any y' in a neighbourhood of y. Conversely, such an estimator gives rise to solutions comprising large strong homogeneity sets. We state that *an estimator $\hat{x}(y)$ is locally strongly homogeneous if and only if the PF φ in* (1-3) *is nonsmooth at* 0. Several PFs nonsmooth at 0 are cited in (4). The corresponding energy $\mathcal{E}_y$ is globally nonsmooth and its minimizers satisfy the necessary condition given in Theorem 1. An equivalent formulation of the energy, as a function of differences $t_k := d_k x$, denoted $\mathcal{F}_y$, is introduced in (7).

To show that nonsmoothness at zero of φ is necessary, we show in Theorem 2 that a minimizer corresponding to a PF which is smooth at zero, *cannot* conserve the strong homogeneity set of any minimizer of $\mathcal{E}_y$ under small data variations.

A PF which is nonsmooth at zero (8) is henceforth used. The necessary conditions for a minimizer $\hat{t}$ of $\mathcal{F}_y$ are given by a system of equations (9-10) and a system of inequalities (11). We establish a sufficient condition for strictness of a minimizer $\hat{\mathbf{t}}$, namely that its smooth part $\hat{\mathbf{t}}_P := \{\,\hat{t}_k;\, k \in \mathcal{S}\setminus\mathcal{J}\}$ is a strict minimum of the restricted energy $\mathcal{F}_y^P$ in (12) and that (11) is a strict (Proposition 1). We also show the local continuity of $\hat{t}_P$ with respect to y (Proposition 2). These constitute the premises for our principal result stated in Theorem 4: if $\hat{t}(y)$ is a strict minimizer for which (11) is strict as well, then there exists an open ball, $B(y;\xi)$, such that for any $y' \in B(y;\xi)$, the set of strong homogeneity remains unchanged, $\mathcal{J}(y) = \mathcal{J}(y')$.

Data space is therefore "partitioned" into volumes which yield the same set of strong homogeneity. This explains why minima of an energy, defined by using a PF which is nonsmooth at zero, tend to exhibit strongly homogeneous zones.

---

1. Introduction

On considère le problème inverse qui consiste à estimer le signal $x \in \mathbb{R}^M$ à partir de l'observation $y = \widetilde{A}x + n$, où $\widetilde{A}$ est un opérateur connu et n un bruit blanc gaussien. La *solution inverse* $\hat{x}$ est définie comme une estimée bayésienne au sens du MAP : c'est le minimiseur d'une énergie *a posteriori* $\mathcal{E}(x) \propto -\ln P(x|y)$ qui relie la log-vraisemblance $-\ln P(y|x)$ et une énergie *a priori* $\Phi(x) \propto -\beta \ln P(x)$ (7). Ainsi, $\hat{x} \equiv \hat{x}(y)$ est une fonction implicite des données $y \in \mathbb{R}^N$:

$$\hat{x} \equiv \hat{x}(y) := \arg\min \mathcal{E}_y(x), \tag{1}$$

$$\mathcal{E}_y(x) := \|\widetilde{A}x - y\|^2 + \beta\Phi(x), \tag{2}$$

$$\Phi(x) := \sum_{k\in\mathcal{S}^\circ} \varphi(d_k x), \quad d_k x := \sum_i d_{k,i} x_i, \tag{3}$$

où d_k sont généralement des opérateurs de différences. On désigne par $\mathcal{S}$ l'ensemble des sites de x et par $\mathcal{S}^\circ \subset \mathcal{S}$ les sites intérieurs – qui sont tels que $d_k x$ ne fait intervenir que des éléments x_i avec $i \in \mathcal{S}$. La *fonction potentiel* (FP) φ est symétrique, croissante sur $[0, \infty[$ et a un minimum strict en 0, elle est C^1 presque partout, et peut être non convexe. Des fonctions $\hat{x}(y)$ de la forme (1-3) émergent dans des contextes disparates, comme l'estimation bayésienne (*voir* [2]) ou la régularisation qu'on applique en estimation (*voir* [11] et [4]) ou en optimisation (*voir* [12] et [8]).

Notre problème est de définir des estimées qui sont *localement fortement homogènes*, c'est-à-dire qui contiennent de larges zones où elles sont localement constantes, ou localement linéaires, *etc.*, et cela en dépit du bruit qui entache les données. Ainsi, demande-t-on que les faibles variations des données laissent inchangées ces zones dans l'estimée. Dans cette Note, nous énonçons qu'*un estimateur de la forme* (1-3) *engendre des solutions comportant des zones fortement homogènes, et conserve ces zones sous de faibles variations des données, si et seulement si la FP* φ *dans* (3) *est non dérivable en zéro.* Parmi les FP les plus utilisées et non dérivables en zéro, on peut citer :

$$\text{(4)} \qquad \begin{array}{l} \text{module : } \varphi(t) = |t|, \qquad \text{concave p.p. : } \varphi(t) = \alpha|t|/(1+\alpha|t|), \\ \text{« 0-1 » : } \varphi(0) = 0\ ,\ \ \varphi(t) = 1 \ \text{ si } \ t \neq 0. \end{array}$$

Afin de simplifier la présentation, *nous supposons que* φ *est continue* C^1 *partout sauf en zéro.*

À notre connaissance, ce problème générique – l'obtention de solutions localement fortement homogènes avec un estimateur de la forme (1-3) – n'a pas été formulé jusqu'à maintenant. Néanmoins, la capacité de la FP *module* d'engendrer des images « presque noires », est analysée à l'aide de la théorie du minimax dans [5]. D'autre part, il est indiqué dans [7] qu'une FP *concave p.p.* produit une solution en marche d'escalier à partir de données en forme de rampe. Notre approche repose sur l'analyse de fonctions non dérivables, qui a été largement développée dans le cadre de l'optimisation (*voir* [10], [1], [12] et [8]).

2. Homogénéité forte et minima relatifs

Les zones « simplement » homogènes correspondent aux positions k où $\boldsymbol{d}_k\hat{\boldsymbol{x}} \approx 0$. Plus fortement :

DÉFINITION 1. – L'*ensemble d'homogénéité forte* $\mathcal{J}$ de $\hat{\boldsymbol{x}}$, relativement aux opérateurs $\boldsymbol{d}_k$, $k \in \mathcal{S}^\circ$, réunit les indices k pour lesquels les différences $\boldsymbol{d}_k\hat{\boldsymbol{x}}$ sont nulles : $\mathcal{J} := \{k \in \mathcal{S}^\circ \ :\ \boldsymbol{d}_k\hat{\boldsymbol{x}} = 0\}$.

Par exemple, si $\boldsymbol{d}_k\hat{\boldsymbol{x}} := \hat{x}_k - \hat{x}_{k-1}$ pour $k \in \mathcal{S}^\circ$, alors $\mathcal{J}$ correspond à des zones localement constantes.

L'ensemble d'homogénéité forte $\mathcal{J} \equiv \mathcal{J}(\boldsymbol{y})$ de $\hat{\boldsymbol{x}}(\boldsymbol{y})$ est une fonction implicite des données $\boldsymbol{y}$. Lorsque φ est convexe, alors $\mathcal{E}_y$ est convexe et possède un minimiseur unique $\hat{\boldsymbol{x}}(\boldsymbol{y})$, donc $\mathcal{J}(\boldsymbol{y})$ est unique. Si φ est non convexe, alors pour $\boldsymbol{y}$ fixé, $\mathcal{E}_y$ présente un ensemble de minimiseurs relatifs (c'est-à-dire locaux) stricts, notés $\hat{\boldsymbol{x}}_n$, $n = 1, \ldots, n_y$. À chacun d'eux correspond un ensemble $\mathcal{J}_n(\boldsymbol{y})$:

$$\text{(5)} \qquad \boldsymbol{y} \ \xrightarrow{\hat{\boldsymbol{x}}}\ \{\hat{\boldsymbol{x}}_1(\boldsymbol{y}), \hat{\boldsymbol{x}}_2(\boldsymbol{y}), \ldots, \hat{\boldsymbol{x}}_{n_y}(\boldsymbol{y})\} \ \xrightarrow{\mathcal{J}}\ \{\mathcal{J}_1(\boldsymbol{y}), \mathcal{J}_2(\boldsymbol{y}), \ldots, \mathcal{J}_{n_y}(\boldsymbol{y})\}.$$

Les énoncés qui suivent concernent *l'homogénéité forte des minimiseurs relatifs stricts* $\hat{\boldsymbol{x}}_n$. Dorénavant, $\hat{\boldsymbol{x}}$ désigne un minimiseur relatif strict quelconque de $\mathcal{E}_y$ et $\mathcal{J}$ son ensemble d'homogénéité forte.

Supposons que $\hat{\boldsymbol{x}} \equiv \hat{\boldsymbol{x}}(\boldsymbol{y})$ contient un ensemble d'homogénéité forte $\mathcal{J}(\boldsymbol{y})$ réaliste. Alors, on souhaite retrouver le même $\mathcal{J}$ à partir de données $\boldsymbol{y}'$ comportant différentes réalisations de bruit.

DÉFINITION 2. – Pour $\boldsymbol{y}$ donné, l'estimateur $\hat{\boldsymbol{x}}(\boldsymbol{y})$ est dit *localement fortement homogène* si l'ensemble $\mathcal{J}(\boldsymbol{y})$ est non vide et reste constant lorsque les données varient à l'intérieur d'une boule ouverte $B(\boldsymbol{y};\varepsilon)$ contenant $\boldsymbol{y}$:

$$\text{il existe } \varepsilon > 0 \ \text{ tel que pour tout } \ \boldsymbol{y}' \in B(\boldsymbol{y};\varepsilon) \ \text{ on ait } \ \mathcal{J}(\boldsymbol{y}) = \mathcal{J}(\boldsymbol{y}').$$

La boule $B(\boldsymbol{y};\varepsilon)$ est définie par rapport à la norme euclidienne $\|.\|$; ainsi, $B(\boldsymbol{y};\varepsilon) := \{\boldsymbol{y}' : \|\boldsymbol{y}'-\boldsymbol{y}\| < \varepsilon\}$. Par la suite, le nombre d'éléments dans $\mathcal{J}$ est noté $\#\{\mathcal{J}\}$.

Lorsque φ est non dérivable en zéro, l'énergie $\mathcal{E}_y$ est non dérivable, ou même discontinue, sur l'union des hyperplans $\cup_{k\in\mathcal{S}^\circ}[\boldsymbol{d}_k\boldsymbol{x} = 0]$.

DÉFINITION 3. – Les dérivées directionnelles gauche et droite de $\mathcal{E}_y$ en $\boldsymbol{x} \in \mathbb{R}^M$ le long d'une direction arbitraire $\boldsymbol{v} \in \mathbb{1}^M$, s'écrivent respectivement $\partial_v^-\mathcal{E}_y(\boldsymbol{x}) := \lim_{h\downarrow 0}[\mathcal{E}_y(\boldsymbol{x} - h\boldsymbol{v}) - \mathcal{E}_y(\boldsymbol{x})]/(-h)$ et $\partial_v^+\mathcal{E}_y(\boldsymbol{x}) := \lim_{h\downarrow 0}[\mathcal{E}_y(\boldsymbol{x} + h\boldsymbol{v}) - \mathcal{E}_y(\boldsymbol{x})]/h$, où $\mathbb{1}^M$ désigne la boule unité dans $\mathbb{R}^M$ par rapport à la norme euclidienne.

Si $\mathcal{E}_y$ est non dérivable en $\boldsymbol{x}$ le long de $\boldsymbol{v}$, alors $\partial_v^+ \mathcal{E}_y(\boldsymbol{x}) \neq \partial_v^- \mathcal{E}_y(\boldsymbol{x})$; elles sont infinies en cas de discontinuité. Les conditions nécessaires de minima relatifs de fonctions continues et non dérivables, sont bien connues (*voir* [12] et [3]), elles s'étendent à certaines fonctions admettant des discontinuités.

THÉORÈME 1. – *Soit $\mathcal{E}_y(\boldsymbol{x})$ continue C^1 par morceaux, comme définie dans* (1-3). *Pour que $\mathcal{E}_y(\boldsymbol{x})$ présente un minimum relatif strict en $\hat{\boldsymbol{x}}$, il faut que*

$$\partial_v^- \mathcal{E}_y(\hat{\boldsymbol{x}}) \leq 0 \leq \partial_v^+ \mathcal{E}_y(\hat{\boldsymbol{x}}) \quad \textit{pour tout} \quad \boldsymbol{v} \in \mathbb{1}^M. \tag{6}$$

Enfin, on exprime $\mathcal{E}_y$ comme une fonction des différences : $t_k = \boldsymbol{d}_k \boldsymbol{x}$ pour $k \in \mathcal{S}^\circ$ et $\boldsymbol{t} = \mathcal{D}\boldsymbol{x}$, où $\mathcal{D}$ est inversible. Posons $\mathcal{A} := \widetilde{\mathcal{A}}\mathcal{D}^{-1} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M]$, où $\boldsymbol{a}_k$ désigne une colonne de $\mathcal{A}$. Alors :

$$\mathcal{F}_y(\boldsymbol{t}) := \mathcal{E}_y(\mathcal{D}^{-1}\boldsymbol{t}) = \|\mathcal{A}\boldsymbol{t} - \boldsymbol{y}\|^2 + \beta \sum_{k \in \mathcal{S}^\circ} \varphi(t_k). \tag{7}$$

Dans cette formulation, l'ensemble d'homogénéité forte est $\mathcal{J}(\boldsymbol{y}) = \{k,\ k \in \mathcal{S}^\circ\ :\ \hat{t}_k(\boldsymbol{y}) = 0\}$.

3. Estimation à l'aide d'une FP dérivable

On s'interroge d'abord sur la capacité d'une FP qui est dérivable en zéro, de conserver un ensemble d'homogénéité forte sous de faibles variations des données.

THÉORÈME 2. – *Soit $\hat{\boldsymbol{t}}(\boldsymbol{y})$ un minimiseur de $\mathcal{F}_y$, tel que* $\#\{\mathcal{J}(\boldsymbol{y})\} > M - \mathrm{rg}(\mathcal{A})$. *Si φ est dérivable en* 0, *alors :*

$$\textit{pour tout} \ \ \varepsilon > 0 \ \ \textit{il existe} \ \ \boldsymbol{y}' \in B(\boldsymbol{y}; \varepsilon) \ \ \textit{tel que} \ \ \mathcal{J}(\boldsymbol{y}) \neq \mathcal{J}(\boldsymbol{y}'),$$

c'est-à-dire qu'il est impossible de conserver l'ensemble d'homogénéité forte $\mathcal{J}(\boldsymbol{y})$ au sens de la définition 2.

Notons que les solutions « intéressantes » contiennent des zones d'homogénéité forte bien larges, c'est-à-dire $\#\{\mathcal{J}\} \gg M - \mathrm{rg}(\mathcal{A})$. On conclut de ce théorème que la non-dérivabilité en zéro de la FP est une *condition nécessaire* pour que $\hat{\boldsymbol{x}}(\boldsymbol{y})$ soit localement fortement homogène.

Exemple. – L'estimateur $\hat{\boldsymbol{t}}(\boldsymbol{y})$, défini à l'aide de $\varphi(t) = t^2$, est explicite ; si de plus $N = M = \mathrm{rg}(\mathcal{A})$, on peut réciproquement exprimer $\boldsymbol{y}$ à partir de $\hat{\boldsymbol{t}}$:

$$\begin{aligned} \hat{\boldsymbol{t}} &= (\mathcal{A}^T\mathcal{A} + \beta\widetilde{I})^{-1}\mathcal{A}^T\boldsymbol{y}, && \text{où} \ \ \widetilde{I}_{k,l} = 1 \text{ si } k = l \in \mathcal{S}^\circ, \ \ \widetilde{I}_{k,l} = 0 \text{ sinon}, \\ \boldsymbol{y} &= \mathcal{C}\hat{\boldsymbol{t}}, && \text{où} \ \ \mathcal{C} := \mathcal{A} + \beta(\mathcal{A}^T)^{-1}\widetilde{I} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M]. \end{aligned}$$

Soit $\hat{t}_k = 0$ pour $k \in \mathcal{J}$. Alors, chaque estimée $\hat{\boldsymbol{t}}(\boldsymbol{y}')$ ayant le même $\mathcal{J}$ s'écrit $\hat{\boldsymbol{t}}(\boldsymbol{y}') = \hat{\boldsymbol{t}} + \boldsymbol{w}$, où $w_k = 0$ si $k \in \mathcal{J}$, et donc elle varie dans un sous-espace de dimension $M - \#\{\mathcal{J}\}$. Comme $\mathcal{C}$ est inversible, $\boldsymbol{y}'$ ne peut varier que dans un sous-espace ayant la même dimension : $\boldsymbol{y}' = \boldsymbol{y} + \sum_{k \in \mathcal{S}\setminus\mathcal{J}} \boldsymbol{c}_k w_k$. Donc, tout $\boldsymbol{y}''$ qui ne se décompose pas de cette manière, mène à $\mathcal{J}(\boldsymbol{y}) \neq \mathcal{J}(\boldsymbol{y}'')$.

4. Minimiseurs correspondant à une FP non dérivable en zéro

Dorénavant, φ est non dérivable en 0, où ses dérivées gauche et droite s'écrivent :

$$\gamma := \varphi'_+(0) = \lim_{h \downarrow 0} \frac{\varphi(h)}{h} = -\varphi'_-(0), \quad \gamma > 0. \tag{8}$$

En particulier, pour la FP « 0-1 », $\gamma = +\infty$.

Le théorème 1 permet d'établir des conditions nécessaires d'un minimum relatif de $\mathcal{F}_y$. Celles-ci généralisent les conditions énoncées dans [1] pour la FP module.

THÉORÈME 3. – *Soit $\hat{t}$ un minimiseur relatif strict de $\mathcal{F}_y$ (7), tel que $\hat{t}_k = 0$ pour $k \in \mathcal{J}$. Alors, $\hat{t}$ satisfait le système :*

$$2a_k^T(y - \mathcal{A}\hat{t}) = 0 \qquad \textit{pour } k \in \mathcal{S} \setminus \mathcal{S}^\circ, \tag{9}$$

$$2a_k^T(y - \mathcal{A}\hat{t}) - \beta\varphi'(\hat{t}_k) = 0 \qquad \textit{pour } k \in \mathcal{S}^\circ \setminus \mathcal{J}, \tag{10}$$

$$2\,|\,a_k^T(y - \mathcal{A}\hat{t})\,| \le \beta\gamma \qquad \textit{pour } k \in \mathcal{J}. \tag{11}$$

Les éléments $\{\,\hat{t}_k,\ k \in \mathcal{S} \setminus \mathcal{J}\,\}$ forment la partie dérivable de $\hat{t}$, tandis que $\{\,\hat{t}_k,\ k \in \mathcal{J}\,\}$ forment sa partie non dérivable (c'est-à-dire nulle). Soient $\mathcal{A}_P$ et $\hat{t}_P$ déduits respectivement de $\mathcal{A}$ et $\hat{t}$ en supprimant les colonnes de $\mathcal{A}$ et les composantes de $\hat{t}$, dont les indices sont dans $\mathcal{J}$: cela donne $\mathcal{A}\hat{t} = \mathcal{A}_P\hat{t}_P$. La restriction de $\mathcal{F}_y$ $\hat{t}_P$, notée $\mathcal{F}_y^P$, est continue $\mathcal{C}^1$ sur un voisinage de $\hat{t}_P$:

$$\mathcal{F}_y^P(\hat{t}_P) = \|\mathcal{A}_P\hat{t}_P - y\|^2 + \beta\sum\nolimits_{k\in\mathcal{S}^\circ\setminus\mathcal{J}}\varphi(\hat{t}_k). \tag{12}$$

La proposition 1 donne une condition *suffisante* pour qu'un minimum de $\mathcal{F}_y$ soit strict.

PROPOSITION 1. – *Soit $\hat{t}$ un minimiseur de $\mathcal{F}_y$ tel que sa partie dérivable $\hat{t}_P$ soit un minimiseur strict de $\mathcal{F}_y^P$ et tel que l'inégalité* (11) *soit stricte. Alors le minimiseur $\hat{t}$ de $\mathcal{F}_y$ est strict.*

D'autre part, $\hat{t}_P(y)$ est une fonction qui est continue localement en y.

PROPOSITION 2. – *Soit $\hat{t}(y)$ un minimiseur strict de $\mathcal{F}_y$ et $\hat{t}_P(y)$ sa partie dérivable* (12). *Alors, il existe deux boules $B(y;\varepsilon)$ et $B(\hat{t}_P;\nu)$, une fonction continue f et une constante C_y, telles que pour tout $y' \in B(y;\varepsilon)$, $\hat{t}'_P := f(y') \in B(\hat{t}_P;\nu)$ est un minimiseur strict de $\mathcal{F}_{y'}^P$ et $\|\hat{t}_P - \hat{t}'_P\| \le C_y\|y - y'\|$.*

5. Condition suffisante d'homogénéité forte

Nous montrons enfin que φ non dérivable en 0 *suffit* pour assurer l'homogénéité forte de $\hat{t}(y)$.

THÉORÈME 4. – *Soit $\hat{t} \equiv \hat{t}(y)$ un minimiseur strict de $\mathcal{F}_y$ pour lequel l'inégalité* (11) *est stricte. Si φ est non dérivable en* 0, *alors*

$$\textit{il existe } \xi > 0 \textit{ tel que pour tout } y' \in B(y;\xi) \textit{ on ait } \mathcal{J}(y) = \mathcal{J}(y'),$$

c'est-à-dire que l'estimateur conserve l'ensemble d'homogénéité forte de la solution lorsque les données varient dans une ξ-boule contenant y.

Un tel minimiseur manifeste une forme de « résistance locale » bien forte : seule sa partie dérivable $\hat{t}_P(y)$ suit les faibles variations des données y, tandis que sa partie non dérivable reste nulle.

Lorsque (11) comprend des inégalités non strictes, l'ensemble $\mathcal{J}$ reste constant seulement pour des données qui varient selon des directions à l'intérieur d'une semi-boule.

PROPOSITION 3. – *Soit $\hat{t}(y)$ tel que* (11) *comprenne plusieurs inégalités non strictes, réunies dans l'ensemble $\widetilde{\mathcal{J}} := \{\,k \in \mathcal{J}\ :\ 2|a_k^T(y - \mathcal{A}\hat{t})| = \beta\gamma\,\}$. On construit la semi-boule :*

$$\mathcal{P}^y := \bigcap\nolimits_{k\in\widetilde{\mathcal{J}}}\{\,u \in \mathbb{1}^N\ :\ \sigma_k\, a_k^T[I - \mathcal{A}_P D_y f(y)]u < 0\,\},\quad \textit{avec } \sigma_k := \operatorname{sign}[a_k^T(y - \mathcal{A}_P\hat{t}_P)],$$

où D désigne un opérateur différentiel. Alors, pour tout $u \in \mathcal{P}^y$, il existe $\tilde{\xi}(u) > 0$ tel que $\mathcal{J}(y + hu) = \mathcal{J}(y)$ pour $0 < h < \tilde{\xi}(u)$.

Évidemment, les variations des données en dehors de $\mathcal{P}^y$ modifient l'ensemble d'homogénéité forte.

Exemple. – On considère l'énergie strictement convexe $\mathcal{E}_y$ définie à l'aide de la FP *module* :

$$\mathcal{E}_y(x) = \|x - y\|^2 + \beta\sum_k |x_k|.$$

L'ensemble d'homogénéité forte s'écrit $\mathcal{J} = \{ k : \hat{x}_k = 0 \}$. Le long d'une direction $\boldsymbol{v}$, on a $[\partial_v^- |x_k|]_{x_k=0} = -|v_k|$ et $[\partial_v^+ |x_k|]_{x_k=0} = |v_k|$. Compte tenu du fait que $\hat{x}_k = 0$ pour $k \in \mathcal{J}$, le minimiseur global $\hat{\boldsymbol{x}}$ est l'unique point qui vérifie les conditions du theoreme 1 :

$$\begin{array}{ll} 2(y_k \quad \hat{x}_k) - \beta \text{ sign } (\hat{x}_k) = 0 & \text{si } k \in \mathcal{S} \setminus \mathcal{J}, \\ 2|y_k| < \beta & \text{si } k \in \mathcal{J}, \end{array} \quad \text{alors} \quad \begin{array}{ll} \hat{x}_k = y_k - \dfrac{\beta}{2} \text{ sign } (y_k) & \text{si } |y_k| > \dfrac{\beta}{2}, \\ \hat{x}_k = 0 & \text{si } |y_k| \leq \dfrac{\beta}{2}, \end{array}$$

La boule $B(\boldsymbol{y}; \xi)$ évoquée au théorème 4 est évidente à construire, pourvu que $y_k \neq \beta/2$, pour tout k. Plus précisément, l'ensemble $\mathcal{J}$ reste inchangé pour tout $|y'_k| < \beta/2$ si $k \in \mathcal{J}$, et pour tout $|y'_k| > \beta/2$ si $k \in$ S \ $\mathcal{J}$. En revanche, si on se trouve dans les conditions de la proposition 3, avec par exemple $y_n = \beta/2$, on a $\mathcal{P}^y = \{ \boldsymbol{u} : u_n < 0 \}$ et $\tilde{\xi} = \beta$.

L'estimateur ci-dessus effectue un *seuillage doux,* d'après la terminologie employee dans [6].

6. L'espace des données et homogénéité forte

Une consequence importante du théorème 4 est que I'espace des données $\mathbb{R}^N$ est « partitionné » en des volumes dont chacun engendre un ensemble d'homogénéité forte. Un tel volume est délimité par le domaine où (11) a lieu et où sa solution correspond effectivement à un minimum relatif de $\mathcal{F}_y$. Notons que les jeux de donnees, pouvant donner lieu à des égalités dans (1 1), sont presque improbables ils sont places sur des hyperplans de $\mathbb{R}^N$, sauf Cventuellement dans des cas très « pathologiques ».

D'autre part, la presence d'une partition en des volumes de $\mathbb{R}^N$, explique pourquoi les minimiseurs d'une Cnergie, comprenant une FP qui est non derivable en zero, ont tendance à former de larges zones d'homogénéité forte.

Les demonstrations des Cnoncts présentés dans cette Note, accompagnées d'amples commentaires et d'illustrations numériques, peuvent Ctre trouvées dans [9].

Remerciements. À Mamadou Mboup pour ses remarques concemant la redaction de cette Note.

Note remise le 17 juin 1997, acceptée le 28 juillet 1997.

Références bibliographiques

[1] **Alliney S. et Ruzinsky** S. **A., 1994.** An algorithm for the minimization of mixed l_1 and l_2 norms with application to Bayesian estimation, *IEEE Transactions on Signal Processing,* SP-42, n° 3, p. 6 18-627.

[2] **Besag J. E., 1989.** Digital image processing : Towards Bayesian image analysis, *J. Appl. Stat., 16,* n° 3, p. 395-407.

[3] **Clarke F., 1989.** *Optimization and Nonsmooth Analysis,* Centre de Recherches Mathtmatiques, Université de Montreal.

[4] **Demoment** G., 1989. Image reconstruction and restoration : Overview of common estimation structure and problems, *IEEE Transactions on Acoustics Speech and Signal Processing,* ASSP-37, n° 12, p. 2024-2036.

[5] **Dohono D., Johnstone I., Hoch J. et Stern A., 1992.** Maximum entropy and the nearly black object, *J. Roy. Stat. Soc. B, 54,* n° 1, p. 41-81.

[6] **Dohono D., 1995.** De-Noising by Soft-Thresholding, *IEEE Transactions on Information Theory, 41,* n° 3, p. 613-627.

[7] **Geman S. et Reynolds** G., 1992. Constrained restoration and recovery of discontinuities, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-14, n° 3, p. 367-383.

[8] **Lemaréchal C. et Sagastidbal** C., 1997. Practical aspects of the Moreau-Yosida regularization I: theoretical properties, *SIAM J. Optimization, 7,* n° 2, p. 367-385.

[9] **Nikolova M., June 1997.** Local strong homogeneity using a regularized estimator, *Rapport interne, UFR Mathématiques et Informatique, Université René-Descartes, 24* p.

[IO] **Shor** *N.* **Z., 1985.** *Minimization methods for Non-differentiable functions,* Springer-Verlag, Berlin, 1e edition.

[I 1] **Titterington D. M., 1985.** Common structure of smoothing techniques in statistics, *Int. Stat. Rev., 53,* n° 2, p. 141-170.

[12] **Hiriart-Urruty J.-B. et Lemaréchal** *C.,* **1996.** *Convex analysis and Minimization Algorithms,* vol. I and II, Springer-Verlag. Berlin.