

Reconstruction de surface par validation croisée

F. Brunet^{1,2,3}, A. Bartoli¹, R. Malgouyres², et N. Navab³

¹ LASMEA UMR 6602, Clermont-Ferrand, France

adrien.bartoli@lasmea.univ-bpclermont.fr

² LAIC EA 2446, Clermont-Ferrand, France

{brunet, malgouyres}@laic.u-clermont1.fr

³ CAMP, Technische Universität München, Allemagne

navab@cs.tum.edu

Mots-Clefs. reconstruction, surface, validation croisée, spline, optimisation

1 Introduction

Nous nous intéressons à la reconstruction d'une surface à partir d'un nuage de points. C'est un problème ayant de nombreuses applications dans des domaines variés tels que l'étude de données statistiques, l'analyse d'images médicales ou encore la CAO. Nous abordons ici plus particulièrement le choix du compromis entre l'attache aux données et la régularité de la surface. L'approche considérée pour résoudre ce problème consiste à minimiser le score de validation croisée. Bien que de nombreux travaux aient déjà été menés à ce sujet, les méthodes utilisées pour optimiser le critère de la validation croisée ne sont que rarement explicitées. De plus les méthodes habituellement décrites sont peu satisfaisantes.

2 La reconstruction de surface

Soit $X = \{(x_i \leftrightarrow y_i) \mid i \in \{1, \dots, n\}, x_i \in \mathbb{R}^N, y_i \in \mathbb{R}^M\}$ un ensemble de données. En pratique, nous considérons principalement le cas $N = 2$ et $M = 1$, c'est-à-dire le cas où y_i représente une élévation par rapport au plan contenant l'ensemble des points x_i . Considérons le modèle de données suivant :

$$\mathbf{y} = f(\mathbf{x}; \mathbf{p}) + \epsilon \quad (1)$$

où $\mathbf{x} = [x_1 \dots x_n]^\top$, $\mathbf{y} = [y_1 \dots y_n]^\top$ et \mathbf{p} est un vecteur donnant les paramètres du modèle. Le problème d'approximation surfacique consiste à trouver le vecteur $\hat{\mathbf{p}}_\lambda$ tel que :

$$\hat{\mathbf{p}}_\lambda = \arg \min_{\mathbf{p}} \mathcal{E}(\mathbf{p}, \lambda) \quad \text{avec} \quad \mathcal{E}(\mathbf{p}, \lambda) = \mathcal{E}_d(\mathbf{p}) + \lambda \mathcal{E}_r(\mathbf{p}) \quad (2)$$

où \mathcal{E}_d est un terme donnant l'énergie d'attache aux données, \mathcal{E}_r quantifie la régularité du modèle et λ contrôle le compromis entre ces deux aspects. Nous ne considérons ici que les modèles linéaires en \mathbf{p} (où A_λ est une matrice dépendant de λ , de \mathbf{x} et du modèle f et où $\tilde{\mathbf{y}}$ contient \mathbf{y}) :

$$A_\lambda \mathbf{p} = \tilde{\mathbf{y}} \quad (3)$$

Les termes d'attache aux données et de régularisation sont souvent choisis comme étant respectivement la moyenne des résidus au carré et l'énergie de torsion [3,5,7] :

$$\mathcal{E}_d(\mathbf{p}) = \frac{1}{n} \sum_{i=1}^n (f(x_i; \mathbf{p}) - y_i)^2 \quad \text{et} \quad \mathcal{E}_r(\mathbf{p}) = \iint_{\mathbb{R}^2} \left\| \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}; \mathbf{p}) \right\|^2 d\mathbf{x} \quad (4)$$

Il a été montré [7] qu'en utilisant de telles énergies, le problème peut être écrit sous la forme de l'équation (3) si l'on utilise comme modèle les *Thin Plate Splines* [1] :

$$f(\mathbf{x} = [\alpha \beta]^\top; \mathbf{p} = [w_1 \dots w_n \ a \ b \ c]^\top) = a\alpha + b\beta + c + \sum_{i=1}^n w_i \rho(x, x_i) \quad (5)$$

$$\text{où } \rho(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|^2 \log \|\mathbf{a} - \mathbf{b}\| \quad (6)$$

Pour un paramètre λ fixé, la solution du problème est donnée par [6,7,4] :

$$\hat{\mathbf{p}}_\lambda = \arg \min_{\mathbf{p}} \|A_\lambda \mathbf{p} - \tilde{\mathbf{y}}\|_2^2 \quad \Leftrightarrow \quad \hat{\mathbf{p}}_\lambda = A_\lambda^\dagger \tilde{\mathbf{y}} \quad (7)$$

Le paramètre λ ne peut évidemment pas être choisi en minimisant directement $\mathcal{E}(\mathbf{p}, \lambda)$. En effet, cette approche conduirait systématiquement à la solution $\lambda = 0$. Nous sélectionnons λ avec une autre technique : la validation croisée [8].

3 La validation croisée et son optimisation

Soit $\hat{\mathbf{p}}_\lambda^{[i]}$ le paramètre minimisant l'énergie du modèle pour λ fixé en considérant le sous-ensemble de données $X^{[i]} = X \setminus \{(x_i \leftrightarrow y_i)\}$. La fonction de validation croisée V est alors donnée par :

$$V(\lambda) = \frac{1}{n} \sum_{i=1}^n \|f(\mathbf{x}; \hat{\mathbf{p}}_\lambda^{[i]}) - y_i\|^2 \quad (8)$$

La sélection automatique du compromis entre l'attache aux données et la régularisation consiste alors à déterminer le paramètre $\hat{\lambda} \in [0, \infty[$ minimisant la fonction V .

Bien qu'expérimentalement l'allure de la fonction V semble être adéquate (V n'est pas convexe mais a un seul minimum, voir figure 1), sa minimisation se heurte à plusieurs problèmes. L'utilisation directe de la formule (8) entrainerait le calcul de n ajustements de surfaces pour une seule valeur de λ ce qui serait bien trop coûteux. Il est possible de montrer [7] que l'expression de V peut être ramenée, de manière équivalente, à :

$$V(\lambda) = \frac{1}{n} \left\| \text{diag} \left(\frac{\mathbf{1}}{\mathbf{1} - \text{diag}(A_\lambda A_\lambda^\dagger)} \right) (A \hat{\mathbf{p}}_\lambda - \mathbf{y}) \right\|^2 \quad (9)$$

La complexité de l'évaluation ponctuelle de V est alors équivalente à celle du calcul d'un seul ajustement. Cependant, cette complexité reste trop élevée pour que le minimum puisse être trivialement déterminé (par exemple, en échantillonnant finement la fonction V). Nous nous intéressons donc à d'autres techniques d'optimisation pour résoudre ce problème. De plus, nous souhaitons prendre en compte des énergies comportant plusieurs termes de régularisation. Plusieurs approches sont considérées : approximation de la fonction de validation croisée, factorisation de la matrice A_λ , utilisation d'autres modèles de surfaces comme les produits tensoriels de B-splines [2].

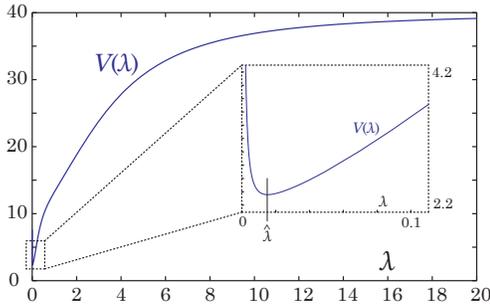


FIG. 1. Fonction de validation croisée.

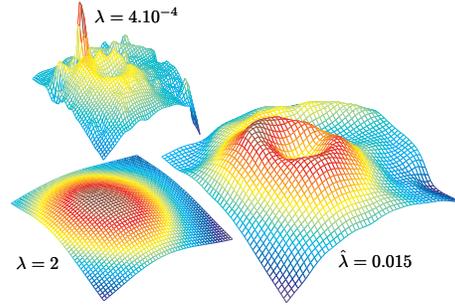


FIG. 2. Ajustement de surface sur le puy Pariou (Auvergne).

Références

1. F. L. Bookstein. Principal warps : Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(6) :567–585, 1989.
2. Paul Dierckx. *Curve and surface fitting with splines*. Oxford University Press, Inc., New York, NY, USA, 1993.
3. G. Donato and S. Belongie. Approximate Thin-Plate Spline mappings. In *ECCV (3)*, pages 21–31, 2002.
4. M. Fornefett, K. Rohr, and H. S. Stiel. Radial Basis Functions with compact support for elastic registration of medical images. *Image and vision computing*, 19 :87–96, 2001.
5. R. Koenker and I. Mizera. Penalized triograms : Total variation regularization for bivariate smoothing. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 66 :145, february 2004.
6. T. Tarpey. A note on the prediction sum of squares statistic for restricted least squares. *The American Statistician*, 54(2) :116–118, May 2000.
7. G. Wahba. *Spline Models for Observational Data*, pages 10–14. SIAM, 1990.
8. G. Wahba and S. Wold. A completely automatic French curve : Fitting spline functions by cross validation. *Communications in Statistics*, 4 :1–18, 1975.