On the Uniqueness of the Solution of Image Reconstruction Problems with Poisson Data

S. Bonettini and V. Ruggiero

Dipartimento di Matematica, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy

Abstract. This paper is concerned with the uniqueness of the Maximum a Posteriori estimate for restoration problems of data corrupted by Poisson noise, when we have to minimize a combination of the generalized Kullback–Leibler divergence and a regularization penalty function. The aim of this paper is to prove the uniqueness result for 2D and 3D problems for several penalty functions, such as an edge preserving functional, a simple case of the class of Markov Random Field (MRF) regularization functionals and the classical Tikhonov regularization.

Keywords: inverse problems, image reconstruction, Poisson noise, edge preserving regularization
PACS: 65R32,65K10,90C90

INTRODUCTION

Poisson data occur in those imaging processes where images are obtained by means of the count of particles, in general photons. In the following we assume that the i–th pixel (voxel) y_i, $i = 1,...,M$ of a 2D (3D) image is an independent realization of a Poisson random variable Y_i. The data y are the detected image of some unknown object $x \in \mathbb{R}^N$ and, in the so–called Bayesian paradigm, x is also assumed to be the realization of a multivalued random variable X. In the case of denoising, the relationship between image and object is simply $Y = X$. In the case of deblurring, if the distortion due to the imaging system can be represented by a linear model, we have $Y = HX + B$. Here $H \in \mathbb{R}^{M\times N}$ is an imaging matrix satisfying the following conditions:

$$H_{lk} \geq 0, \quad \sum_{k=1}^{N} H_{lk} > 0, \quad l = 1,...,M, \quad \sum_{l=1}^{M} H_{lk} > 0, \quad k = 1,...,N. \tag{1}$$

Moreover, B is a random variable with Poisson distribution with expected value be_M, $b > 0$, representing the background emission. Here and in the following e_M is the vector in $\mathbb{R}^M$ with all entries equal to one. According to the previous assumptions, for a given y, the maximum likelihood (ML) estimate [9] of the expected value of the object can be obtained by minimizing the following generalized Kullback–Leibler divergence

$$f_0(x) = \sum_{k\in\mathscr{I}_1} \left\{ y_k \ln \frac{y_k}{(Hx+be_M)_k} + (Hx+be_M)_k - y_k \right\} + \sum_{k\in\mathscr{I}_2} (Hx+be_M)_k \tag{2}$$

where $\mathscr{I}_1 = \{k \in \{1,...,N\} : y_k > 0\}$ and $\mathscr{I}_2 = \{k \in \{1,...,N\} : y_k = 0\}$. In general, $\mathscr{I}_1 \neq \emptyset$. The denoising case is recovered by setting $H = I$ and $b = 0$. The function $f_0(x)$ is nonnegative, convex and coercive [1]. *A priori* information is introduced in the Bayesian approach [6] by assuming a probability distribution also on the random variable X. A very frequent assumption is a Gibbs distribution, usually called *prior*, $p_X(x) = \frac{1}{K}e^{-\beta f_1(x)}$, where K is a normalization constant and $f_1(x)$ is a given penalty function. Once the prior has been chosen, considering that in the image reconstruction problems only nonnegative solutions have physical meaning, the *maximum a posteriori* (MAP) estimate of the expected value of the object can be determined by solving the following constrained optimization problem

$$\begin{array}{ll} \min & f_\beta(x) = f_0(x) + \beta f_1(x) \\ \text{subject to} & x \in \Omega \end{array} \tag{3}$$

where β is a positive regularization parameter, and Ω is a closed convex subset of the nonnegative orthant of $\mathbb{R}^N$.
The aim of this note is to prove the uniqueness of the solution of problem (3) when x is a 2D or 3D object, and $f_0(x)$ is

CP1281, *ICNAAM, Numerical Analysis and Applied Mathematics, International Conference 2010*
edited by T. E. Simos, G. Psihoyios, and Ch. Tsitouras

the Kullback–Leibler divergence (2) (see [1] for the 1D case). In particular, we consider the following – nonnegative, differentiable and convex – penalty functions:

- a regularization function from the class of the edge preserving potentials [5],
- a simple case of the class of the Markov Random Field (MRF) regularization functionals [7],
- the classical Tikhonov regularization.

We prove that $f_\beta(x)$ is strictly convex by showing that the intersection between the null spaces of the Hessian matrices of f_0 and f_1, i.e. the null space of $\nabla^2 f_\beta(x)$, is trivial. In particular, we show that the the null space of the Hessian of the penalty function $f_1(x)$ is the set

$$\mathscr{N}[\nabla^2 f_1(x)] = \{\alpha e_N : \alpha \in \mathbb{R}\} \;\; \forall x \in \mathbb{R}^N \tag{4}$$

and, furthermore, that e_N does not belong to the null space of $f_0(x)$. The strict convexity of $f_\beta(x)$ is crucial in the convergence analysis of the numerical optimization methods for the solution of (3) and in the theory of a recently proposed discrepancy principle for Poisson data [1].

NOTATION AND PRELIMINARY RESULTS

We denote the ℓ_2 norm of a vector $x \in \mathbb{R}^N$ as $\|x\|$. In the restoration problems, we indicate the entries of the unknown object $x \in \mathbb{R}^N$ as x_k, $k = 1, \ldots, N$ where, in the 2D and 3D cases, k corresponds to a multi-index. In particular, assuming a column–wise ordering, we define $x_k \equiv x_{ij}$, $k = (j-1)n + i$ for a 2D image $x \in \mathbb{R}^{n\times m}$ $(N = nm)$ and $x_k \equiv x_{ijh}$, $k = (h-1)nm + (j-1)n + i$ for a 3D image $x \in \mathbb{R}^{n\times m\times r}$ $(N = nmr)$. In the following we employ both notations, with a single index or a multi-index, taking account of the previous equivalences. We consider two kinds of boundary conditions:

P) periodic boundary conditions; for example, in the 2D case, we have: $x_{n+1j} = x_{1j}$, $x_{0j} = x_{nj}$, $j = 1, \ldots, m$ and $x_{im+1} = x_{i1}$, $x_{i0} = x_{im}$, $i = 1, \ldots, n$; we define also $x_{00} = x_{nm}$, $x_{n+1m+1} = x_{11}$, $x_{n+10} = x_{1m}$ and $x_{0m+1} = x_{n1}$;

R) reflexive (or Neumann) boundary conditions; this means, in the 2D case, that we assume $x_{n+1j} = x_{nj}$, $x_{0j} = x_{1j}$, $x_{in+1} = x_{in}$ and $x_{i0} = x_{i1}$, for $i = 1, \ldots, n$, $j = 1, \ldots, m$; we define also $x_{00} = x_{11}$, $x_{n+1m+1} = x_{nm}$, $x_{n+10} = x_{n1}$ and $x_{0m+1} = x_{1m}$.

The gradient and the Hessian matrix of $f_0(x)$ are given by

$$\nabla f_0(x) \;=\; e_N - H^T Z^{-1} y \quad , \quad \nabla^2 f_0(x) \;=\; H^T Y Z^{-2} H$$

where Z and Y are diagonal matrices of order M, such that $(Z)_{kk} = (Hx + be_M)_k$, $k \in \mathscr{I}_1$, and $(Z)_{kk} = 1$, $k \in \mathscr{I}_2$ and $(Y)_{kk} = y_k$, $k = 1, \ldots, M$. Thus, the null space of $\nabla^2 f_0(x)$ is

$$\mathscr{N}[\nabla^2 f_0(x)] = \{u \in \mathbb{R}^N : (Hu)_k = 0 \text{ if } k \in \mathscr{I}_1\}$$

and it contains the null space of H. Under the assumption (1), we remark that $e_N \notin \mathscr{N}[\nabla^2 f_0(x)]$ for both denoising $(H = I,\, b = 0)$ and deblurring problems. As a consequence, if property (4) holds for the penalty function $f_1(x)$, we have $\mathscr{N}[\nabla^2 f_\beta(x)] = \{0\}$. In the following sections we will show that (4) holds for several kinds of convex penalty functions in the 1D, 2D and 3D cases, yielding the strict convexity of $f_\beta(x)$.

EDGE PRESERVING REGULARIZATION

Hypersurface and Total Variation regularization

This section is focused on the edge preserving penalty function having the following form

$$f_1(x) = \sum_{k=1}^{N_1} \sqrt{D_k} \tag{5}$$

where $D_k = (x_{k+1} - x_k)^2 + \delta^2$, $D_k \equiv D_{ij} = (x_{i+1j} - x_{ij})^2 + (x_{ij+1} - x_{ij})^2 + \delta^2$ or $D_k \equiv D_{ijh} = (x_{i+1jh} - x_{ijh})^2 + (x_{ij+1h} - x_{ijh})^2 + (x_{ijh+1} - x_{ijh})^2 + \delta^2$ in the 1D, 2D and 3D case, respectively. Here δ is a nonzero scalar parameter while $N_1 = N$ for periodic boundary conditions and $N_1 = n - 1$, $N_1 = nm - 1$ or $N_1 = (nm-1)r$ for reflexive boundary conditions in the 1D, 2D and 3D case, respectively.

The model function in (5), for small values of δ, can be considered as a discrete approximation of the Total Variation functional [10]. Moreover, it formally describes also the hypersurface potential proposed in [5].

Following [4], we can also write $D_k = \|A_k x\|^2 + \delta^2$ where A_k, $k = 1,\dots,N_1$, is a matrix with N columns and $1,2$ or 3 rows, depending on the dimension of the object x and on the assumed boundary conditions. In particular, any A_k has exactly two nonzero elements for each row, which are equal to 1 and -1. For further details, see [2].

With these settings, the function (5) can be written also as a sum of norms. Now, in order to derive a matrix expression for the gradient and Hessian of (5), which is formally independent on the dimension of the object and on the boundary conditions, we define the block matrices

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \dots \\ A_{N_1} \end{pmatrix}, \qquad E(x) = \mathrm{diag}(D_k^{1/2} I_{d_k})_{k=1,\dots,N_1}, \qquad F(x) = \mathrm{diag}\left(I_{d_k} - \frac{1}{D_k} A_k x x^T {A_k}^T\right)_{k=1,\dots,N_1}. \tag{6}$$

The matrix A is a $q \times N$ matrix: for periodic boundary conditions $q = dN_1$, where $d = 1,2,3$ is the dimension of the object x, while for reflexive boundary conditions $q = n-1$, $q = 2(n-1)(m-1)+n+m-2$ or $q = 3(n-1)(m-1)(r-1)+2(n-1)(m-1)+2(r-1)(n-1)+2(r-1)(m-1)+n+m+r-3$ for the 1D, 2D and 3D case, respectively.

The matrices $E(x), F(x) \in \mathbb{R}^{q\times q}$ are block diagonal matrices where the dimension d_k of the k–th diagonal block is the number of rows of the difference matrix A_k ($d_k = d$ for all k when periodic boundary conditions are assumed). Then, the gradient and Hessian of $f_1(x)$ can be written as

$$\nabla f_1(x) = A^T E(x)^{-1} A x, \qquad \nabla^2 f_1(x) = A^T E(x)^{-1} F(x) A.$$

In particular, $E(x)$ is diagonal with positive diagonal entries and the k–th diagonal block of $F(x)$ is the difference between the identity matrix and a dyadic product; from the definition of D_k, we have that $\frac{1}{D_k} x^T {A_k}^T A_k x = \frac{\|A_k x\|^2}{D_k} < 1$. Consequently, from the Sherman-Morrison theorem (see for example [8, p. 19]), it follows that all the diagonal blocks of $F(x)$ and, thus, $F(x)$ itself, are nonsingular. Taking account of the previous remarks, we can characterize the null space of $\nabla^2 f_1(x)$ in the following way.

Theorem 1 *The null space of the Hessian matrix $\nabla^2 f_1(x)$ is given by the set of the minimum points of $f_1(x)$.*

Proof. Since $E(x)^{-1}F(x)$ is nonsingular for all $x \in \mathbb{R}^N$, we have that $\mathscr{N}[\nabla^2 f_1(x)] = \mathscr{N}[A]$. On the other hand, the minimum points of the convex function $f_1(x)$ satisfy the stationarity condition $\nabla f_1(x) = 0$, that is $A^T E(x)^{-1} A x = 0$. Then, the set of the minimum points of $f_1(x)$ is the subspace $\mathscr{N}[A]$. □

Since the minimum points of the functional (5) in both cases of periodic or reflexive boundary conditions are all the points $x \in \mathbb{R}^N$ such that $x_1 = x_2 = \dots = x_N$, the previous theorem implies that the property (4) holds in the 1D, 2D and 3D cases.

Markov Random Field regularization

For the 2D edge preserving image reconstruction, we consider the following penalty function [7]

$$f_1(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\ell \in \mathscr{N}_{ij}} \sqrt{(x_{ij} - x_{ij}^{(\ell)})^2 / (w_{ij}^{(\ell)})^2 + \delta^2} \tag{7}$$

where $x_{ij}^{(\ell)}$, $\ell \in \mathscr{N}_{ij} \subseteq \{1,\dots,8\}$ is the set of indices of the first neighbors of the pixel ij, while $w_{ij}^{(\ell)}$ are positive weights. The function (7) is a simple example of the Markov random field regularization [6], where $f_1(x)$ represents a sum of potentials involving only the "nearest neighbor" interactions between the components of x.

When we assume periodic boundary conditions, $\mathscr{N}_{ij} = \{1,\dots,8\}$ for all $i = 1,\dots,n$, $j = 1,\dots,m$. Then, for any $k = 1,\dots,nm$, we can consider eight $1 \times nm$ vectors $A_k^{(1)},\dots,A_k^{(8)}$, with all components equal to zero except two entries, which are equal to $w_k^{(\ell)}$ and $-w_k^{(\ell)}$, corresponding to the indices of x_{ij} and $x_{ij}^{(\ell)}$, respectively. We now define the following square matrices of order nm

$$A^{(\ell)} = \begin{pmatrix} A_1^{(\ell)} \\ \vdots \\ A_{nm}^{(\ell)} \end{pmatrix}, \qquad E(x)^{(\ell)} = \mathrm{diag}\left({D_{ij}^{(\ell)}}^{\frac{1}{2}}\right), \quad \ell = 1,\dots,8,$$

where $D_k^{(\ell)} \equiv D_{ij}^{(\ell)} = ((x_{ij} - x_{ij}^{(\ell)})/w_{ij}^{(\ell)})^2 + \delta^2$.

When reflexive boundary conditions are assumed, we have that $\mathscr{N}_{ij} = \{1,...,8\}$ for $i = 2,...,n-1$, $j = 2,...,m-1$, while, for the corners $((i,j) = (1,1),(n,1),(1,m),(n,m))$, $\mathscr{N}_{ij}$ contains only 3 indices, and for the remaining boundary pixels, $\mathscr{N}_{ij}$ has only 5 indices. Then, the number $N_1^{(\ell)}$ of rows of the matrices $A^{(\ell)}$ and $E(x)^{(\ell)}$ depends on the value of ℓ. In any case, with both periodic and reflexive boundary conditions, the gradient and Hessian of (7) have the following form

$$\nabla f_1(x) = \left(\sum_{\ell=1}^{8} A^{(\ell)^T} E(x)^{(\ell)^{-1}} A^{(\ell)} \right) x, \qquad \nabla^2 f_1(x) = \sum_{\ell=1}^{8} A^{(\ell)^T} E(x)^{(\ell)^{-1}} F(x)^{(\ell)} A^{(\ell)}$$

where $F(x)^{(\ell)} = \mathrm{diag}(\delta^2/D_k^{(\ell)})_{k=1,\dots,N_1}$ is a diagonal matrix with positive entries. The Hessian of the function (7) is the sum of the eight positive semidefinite matrices $A^{(\ell)^T} E(x)^{(\ell)^{-1}} F(x)^{(\ell)} A^{(\ell)}$; then, we have $\mathscr{N}[\nabla^2 f_1(x)] = \bigcap_{\ell=1}^{8} \mathscr{N}[A^{(\ell)}]$ for all $x \in \mathbb{R}^{nm}$. As in Theorem 1, we can conclude that the null space of the Hessian matrix is the set of the minimum points of (7), that is $\{\alpha e_N : \alpha \in \mathbb{R}\} = \mathscr{N}[A^{(\ell)}] = \bigcap_{\ell=1}^{8} \mathscr{N}[A^{(\ell)}] = \mathscr{N}[\nabla^2 f_1(x)]$. Similar conclusions can be extended to 3D case.

Tikhonov regularization

We consider the standard Tikhonov regularization, based on the ℓ_2 norm of the discrete gradient of x

$$f_1(x) = \frac{1}{2} \sum_{k=1}^{N_1} D_k = \frac{1}{2} \sum_{k=1}^{N_1} \|A_k x\|^2 \tag{8}$$

where D_k, A_k and N_1 are defined as in section on hypersurface. The regularization function $f_1(x)$ is a convex quadratic function and its gradient and Hessian are given by

$$\nabla f_1(x) = A^T A x, \qquad \nabla^2 f_1(x) = A^T A$$

where A is given in (6). Under periodic or reflexive boundary condition, property (4) holds for the Tikhonov regularization function (8) for $d = 1,2,3$. Indeed, we have $\mathscr{N}[\nabla^2 f_1(x)] = \mathscr{N}[A]$ and, following the proof of Theorem 1, we can prove property (4) also for the function (8).

APPLICATIONS

The uniqueness result proved in the previous sections, can be applied to the convergence analysis of the Scaled Gradient Projection (SGP) method proposed in [3]. Indeed, the strict convexity of the objective function, which is also coercive, for the penalty functions (5), (7) and (8), implies the convergence of the sequence generated by the algorithm SGP to the unique solution of (3). Indeed, this is a direct consequence of Theorem 2.1 in [3].

Furthermore, the uniqueness of the solution of (3), besides the property $\nabla f_1(x)^T e_N = 0$ which can be easily verified by the definitions of $\nabla f_1(x)$, is crucial for the well posedness of the *discrepancy principle* for Poisson data recently proposed in [1].

REFERENCES

1. M. Bertero, P. Boccacci, G. Talenti, R. Zanella and L. Zanni, *Inverse Problems*, to appear, (2010).
2. S. Bonettini, V. Ruggiero, http://eprints.unife.it/249/ (2010).
3. S. Bonettini, R. Zanella and L. Zanni, *Inverse Problems* **25**, 015002 (2009).
4. T.F. Chan, G.H. Golub and P. Mulet, *SIAM J. Sci. Comput.* **20**, 1964–1977 (1999).
5. P. Charbonnier, L. Blanc–Féraud, G. Aubert and A. Barlaud, *IEEE Trans. Image Processing* **6**, 298–311 (1997)
6. S. Geman and D. Geman, *IEEE Trans. Pattern Anal. Intell.* **6**, 721–741 (1984).
7. S. Geman, K. Manbeck and D.E. McClure "A comprehensive statistical model for single photon emission tomography" in *Markov Random Fields: theory and Applications* edited by R. Chellappa et al., Academic Press, Boston, 1993, pp. 93–130.
8. R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.
9. L.A. Shepp and Y. Vardi, *Trans. Med. Imaging* **MI-1**, 113–122 (1982)
10. C.R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.