J. Inv. Ill-Posed Problems **15** (2007), 12–8

DOI 10.1515 / JIP.2007.002

A Theoretical Framework for the Regularization of Poisson Likelihood Estimation Problems

Johnathan M. Bardsley

Communicated by Name of Editor

Abstract. Let $z = Au + \gamma$, where $\gamma > 0$ is constant, be an ill-posed, linear operator equation. Such a model arises, for example, in both astronomical and medical imaging, in which case γ corresponds to background light intensity. Regularized solutions of this equation can be obtained by solving

$$R_\alpha(A, z) = \arg\min_{u \geq 0}\{T_0(Au; z) + \alpha J(u)\},$$

where $T_0(Au; z)$ is the negative-log of the Poisson likelihood functional, and $\alpha > 0$ and J are the regularization parameter and functional, respectively. This variational problem can be motivated from the fact that typical image data contains Poisson noise, and it has been analyzed, for three different choices of J, in previous work of the author. Our goal in this paper is to prove that these previous results imply that the approach defines a *regularization scheme*—which we rigorously define here—for each choice of J. Determining the appropriate definition for *regularization scheme* in this context is important: not only will it serve to unify the previously mentioned theoretical arguments, it will provide a framework for future theoretical analysis. In addition, we modify our presentation somewhat in order to improve understandability and provide missing arguments from our previous analysis.

Key words. regularization, Poisson likelihood, statistical estimation, mathematical imaging.

AMS classification. 65J22, 65K10, 65F22.

1. Introduction

Consider the linear operator equation

$$z = Au + \gamma. \tag{1.1}$$

Our application of interest is image processing, so that $z \in L^\infty(\Omega)$ denotes the image intensity and $u \in L^2(\Omega)$ the intensity of the unknown object; both are defined on a closed, bounded domain $\Omega \subset \mathbb{R}^d$. The positive constant γ corresponds to the intensity of the background, which is typically due to light sources outside of the field of view of the imaging instrument and is standard in many imagining models. Finally,

$$Au(x) \stackrel{\text{def}}{=} \int_\Omega a(x; y)u(y)\, dy,$$

where $a \in L^\infty(\Omega \times \Omega)$ is the error free, nonnegative point spread function. Given these assumptions, $A : L^2(\Omega) \to L^2(\Omega)$ is a compact operator, and hence, the problem of

This work was supported by the NSF under DMS-???.

solving (1.1) for u is ill-posed [8, 9]. Moreover, $Au \geq 0$ whenever $u \geq 0$, and hence, assuming that the true image $u_{\text{exact}} \geq 0$, we have that the error free data $z = Au_{\text{exact}} + \gamma$ is bounded below by γ.

In previous works of the author [2, 3, 4], the following variational problem is theoretically analyzed:

$$\arg\min_{u \in \mathcal{C}} \left\{ T_\alpha(Au; z) \stackrel{\text{def}}{=} T_0(Au; z) + \alpha J(u) \right\}, \tag{1.2}$$

where

$$\mathcal{C} = \{u \in L^2(\Omega) \mid u \geq 0\},$$

α and J are the regularization parameter and functional, respectively, and

$$T_0(Au; z) \stackrel{\text{def}}{=} \int_\Omega ((Au + \gamma) - z \log(Au + \gamma))\, dx. \tag{1.3}$$

T_0 is the functional analogue of the negative-log of the Poisson likelihood function [4, 5], which arises in image deblurring when a charge-coupled-device (CCD) camera is used to collect images [7].

The definition for T_0 used in [2, 3, 4] is given by $T_0(Au + \sigma^2; z + \gamma + \sigma^2)$, with T_0 defined by (1.3). This modification is motivated by the fact that we are using (1.1) instead of $z = Au$ as our model and that we have removed the parameter σ^2 due to the fact that it stems from a noise model for CCD camera data [7] and, as such, is extraneous in the functional (non-stochastic and non-discrete) setting of this paper.

Returning to (1.2), we note that in [2], the regularization functional

$$J(u) = \|u\|^2_{L^2(\Omega)}, \tag{1.4}$$

where $\|u\|^2_{L^2(\Omega)} = \int_\Omega u^2\, dx$, is considered; whereas in [3], the regularization functional

$$J(u) = \|\mathbf{\Lambda}(x)\nabla u\|^2_{L^2(\Omega)}, \tag{1.5}$$

where $\mathbf{\Lambda}(x)$ is a 2×2 positive definite matrix with continuously differentiable components for all $x \in \Omega$, is the focus; and finally, in [4], we consider the total variation regularization functional

$$J(u) = J_\beta(u) \stackrel{\text{def}}{=} \sup_{v \in \mathcal{V}} \int_\Omega \left(-u\nabla \cdot v + \sqrt{\beta(1 - |v(x)|^2)} \right) dx, \tag{1.6}$$

where $\beta \geq 0$ and

$$\mathcal{V} = \{v \in C_0^1(\Omega; \mathbb{R}^d) \ : \ |v(x)| \leq 1 \quad x \in \Omega\}.$$

We note that if u is continuously differentiable on Ω, (1.6) takes the recognizable form [1, Theorem 2.1]

$$J_\beta(u) \stackrel{\text{def}}{=} \int_\Omega \sqrt{|\nabla u|^2 + \beta}\, dx. \tag{1.7}$$

$J_0(u)$ is known as the *total variation* of u.

Each of the above regularization functions is convex. Moreover, the Tikhonov functional is strongly convex. In order to see that T_0 is convex, we note that the gradient and Hessian of T_0 are given, respectively, by

$$\nabla T_0(Au;z) = A^*\left(\frac{Au-(z-\gamma)}{Au+\gamma}\right), \tag{1.8}$$

$$\nabla^2 T_0(Au;z) = A^*\text{diag}\left(\frac{z}{(Au+\gamma)^2}\right)A, \tag{1.9}$$

where "*" denotes operator adjoint. Given our assumptions regarding z and A we see immediately that the Hessian is positive semi-definite for all $u \in \mathcal{C}$ and is positive definite provided A is an invertible operator. Thus we have that T_α is convex in all cases and is strictly convex when J is given by (1.4) or when A is an invertible operator. Moreover, we see that $\nabla T_0(Au_{\text{exact}};z) = 0$, and hence, u_{exact} is a nonnegatively constrained minimizer of T_0 that is unique if A is an invertible operator.

The goal of this paper is to present a theoretical framework that unifies the analysis of (1.2) found in [2, 3, 4], and in turn, to provide a framework within which future theoretical arguments can be made. To this end, we define the term *regularization scheme* in Section 2, and then in Section 3 prove, citing the results of [2, 3, 4], that problem (1.2) defines a *regularization scheme*.

2. Definition of Regularization Scheme

The classical theory of regularization [9, 8] requires that we prove that our regularization method (1.2) is a *regularization scheme* for (1.1) for each of the above functionals J. As we will see, the results found in [2, 3, 4] imply this, however in those papers this term was not defined. We do that in this section.

First, we define a sequence of operator equations

$$z_n = A_n u + \gamma, \tag{2.1}$$

where $z_n \in L^\infty(\Omega)$ is nonnegative, $u \in L^2(\Omega)$, $\gamma > 0$, and

$$A_n u(x) \stackrel{\text{def}}{=} \int_\Omega a_n(x;y)u(y)\,dy,$$

with $a_n \in L^\infty(\Omega \times \Omega)$ a nonnegative point spread function. Note, then, that $A_n u \geq 0$ whenever $u \in \mathcal{C}$.

Next, we define

$$\mathcal{B} = \{\hat{a} \in L^\infty(\Omega\times\Omega) \mid \hat{a} \geq 0\} \times \{\hat{z} \in L^\infty(\Omega) \mid \hat{z} \geq 0\},$$

which is clearly a closed subspace of the Banach space $L^\infty(\Omega\times\Omega)\times L^\infty(\Omega)$, whose norm we denote $\|\cdot\|_{\mathcal{B}}$.

Finally, we define the operator $R_\alpha : \mathcal{B} \to \mathcal{C}$ by

$$R_\alpha(a, z) \stackrel{\text{def}}{=} \arg\min_{u \in \mathcal{C}} T_\alpha(Au; z), \tag{2.2}$$

and we are ready to present our definition.

Definition 2.1. $\{R_\alpha\}_{\alpha>0}$ defines a regularization scheme on $\mathcal{B}$ provided
1. R_α is well-defined and continuous on $\mathcal{B}$ for all $\alpha > 0$;
2. given a sequence $\{(a_n, z_n)\}_{n=1}^\infty \subset \mathcal{B}$ such that $\|(a_n, z_n) - (a, z)\|_{\mathcal{B}} \to 0$, there exists a positive sequence $\{\alpha_n\}_{n=1}^\infty$ such that

$$\|R_{\alpha_n}(a_n, z_n) - u_{\text{exact}}\|_{L^p(\Omega)} \to 0$$

for some $p \geq 1$.

In the next section, we show that $R_\alpha(a, z)$ satisfies conditions 1 and 2 of Definition 2.1.

3. $R_\alpha(a, z)$ Defines a Regularization Scheme

3.1. $R_\alpha(a, z)$ is a Well-Defined Operator on $\mathcal{B}$

We begin by defining two function spaces that will be of import in the discussion that follows. First, we define

$$\langle u, v\rangle_{H^1(\Omega)} \stackrel{\text{def}}{=} \langle u, v\rangle_{L^2(\Omega)} + \langle \nabla u, \nabla v\rangle_{L^2(\Omega)}, \tag{3.1}$$

where " ∇ " denotes the gradient. The set of all functions $u \in C^1(\Omega)$ such that

$$\|u\|_{H^1(\Omega)} = \sqrt{\langle u, u\rangle_{H^1(\Omega)}}$$

is finite is a normed linear space whose closure in $L^2(\Omega)$ is the Sobolev space $H^1(\Omega)$ [6]. We note, moreover, that with the inner-product defined in (3.1), $H^1(\Omega)$ is a Hilbert space.

Next, the space of bounded variation is defined

$$BV(\Omega) = \{u \in L^1(\Omega) : J_0(u) < +\infty\}, \tag{3.2}$$

where J_0 is defined by (1.6). $BV(\Omega)$ is a Banach space with norm

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + J_0(u).$$

We now prove that solutions of (1.2) exist and are unique under certain reasonable assumptions. Recall from our discussion in the Introduction that T_α is convex in all cases, and is strictly convex provided J is given by (1.4) or when A is an invertible operator. However, in order to prove that solutions of (1.2) exist, we also need that T_α is coercive, which is defined

$$T_\alpha(Au; z) \to +\infty \quad \text{whenever} \quad \|u\| \to +\infty.$$

Here $\|\cdot\|$ is a norm that depends upon the choice of J: for J given by (1.4), $\|\cdot\| \stackrel{\text{def}}{=} \|\cdot\|_{L^2(\Omega)}$; for J given by (1.5), $\|\cdot\| \stackrel{\text{def}}{=} \|\cdot\|_{H^1(\Omega)}$; and for J given by (1.6), $\|\cdot\| \stackrel{\text{def}}{=} \|\cdot\|_{BV(\Omega)}$. Proofs of the coercivity of T_α for each case can be found in [2, 3, 4]. It follows, then, that solutions of (1.2) exist and are unique—making $R_\alpha(a, z)$ a well-defined operator—when J is given by (1.4) or when A is invertible. We will make these assumptions throughout the remainder of the paper.

Another question of interest is, what is the range of R_α? The proofs of existence and uniqueness of minimizers of T_α found in [2, 3, 4] indicate that when J is given by (1.4), $\text{Range}(R_\alpha) \subset L^2(\Omega)$, when J is given by (1.5), $\text{Range}(R_\alpha) \subset H^1(\Omega)$, and when J is given by (1.6), $\text{Range}(R_\alpha) \subset BV(\Omega)$. Thus the choice of regularization functional has a significant effect on the properties of the regularized solution, as is readily verified by numerical experiment (see [2, 3, 4]).

3.2. $R_\alpha(a, z)$ is a Continuous Operator on $\mathcal{B}$

For each choice of regularization functional J, the theoretical arguments in [2, 3, 4] give us that $\|R_\alpha(a_n, z_n) - R_\alpha(a, z)\|_{L^p(\Omega)} \to 0$, for some $p \geq 1$, provided $\|A_n - A\|_{L^1(\Omega)}, \|z - z_n\|_{L^\infty(\Omega)} \to 0$. Noting that

$$\begin{aligned} \|(\hat{A} - A)u\|_{L^1(\Omega)} &= \int_\Omega \left| \int_\Omega (\hat{a}(x, y) - a(x, y))u(y)\, dy \right| dx \\ &\leq \|\hat{a} - a\|_{L^\infty(\Omega \times \Omega)} \cdot \|u\|_{L^1(\Omega)}, \end{aligned}$$

we see that

$$\|\hat{A} - A\|_{L^1(\Omega)} = \max_{u \in L^1(\Omega)} \frac{\|(\hat{A} - A)u\|_{L^1(\Omega)}}{\|u\|_{L^1(\Omega)}} \leq \|\hat{a} - a\|_{L^\infty(\Omega \times \Omega)},$$

and hence, $\|\hat{a} - a\|_{L^\infty(\Omega \times \Omega)} \to 0$ implies $\|\hat{A} - A\|_{L^1(\Omega)} \to 0$. Thus we have that if $\|(a_n, z_n) - (a, z)\|_{\mathcal{B}} \to 0$ then $\|R_\alpha(a_n, z_n) - R_\alpha(a, z)\|_{L^p(\Omega)} \to 0$, where $p = 2$ when J is given by (1.4) or (1.5), and $1 \leq p < d/(d-1)$ when J is given by (1.6), where d is the dimension of the computational domain.

However, we have found an error in the continuity arguments found in [2]. In particular, these arguments only give the result that $R_\alpha(a_n, z_n)$ converges to $R_\alpha(a, z)$ weakly in $L^2(\Omega)$, but strong convergence in $L^2(\Omega)$ follows if $||R_\alpha(a_n, z_n)||_{L^2(\Omega)} \to ||R_\alpha(a, z)||_{L^2(\Omega)}$. We prove this now. By the weak lower semi-continuity of the $L^2(\Omega)$ norm $||R_\alpha(a, z)||_{L^2(\Omega)} \leq \liminf ||R_\alpha(a_n, z_n)||_{L^2(\Omega)}$. However

$$T_\alpha(A_n(R_\alpha(a_n, z_n)), z_n) \leq T_\alpha(A_n(R_\alpha(a, z)), z_n)$$

implies

$$\begin{aligned}\liminf ||R_\alpha(a_n,z_n)||^2_{L^2(\Omega)} &\leq ||R_\alpha(a,z)||^2_{L^2(\Omega)} + 2\alpha^{-1}[T_0(A_n(R_\alpha(a,z))) \\ &\qquad - \liminf T_0(A_n(R_\alpha(a_n,z_n)),z_n))] \\ &\leq ||R_\alpha(a,z)||^2_{L^2(\Omega)},\end{aligned}$$

since T_0 is weakly, lower semi-continuous. Hence $||R_\alpha(a_n,z_n)||_{L^2(\Omega)}$ converges strongly to $||R_\alpha(a,z)||_{L^2(\Omega)}$, which gives us the result.

Thus R_α satisfies condition 1 of Definition 2.1 for J given by (1.4), (1.5), and (1.6).

3.3. $R_\alpha(a,z)$ is Convergent

Finally, we must show that condition 2 of Definition 2.1 holds. First, we note that by our discussion above, if $\|(a_n,z_n)-(a,z)\|_{\mathcal{B}} \to 0$ then $\|A_n - A\|_{L^1(\Omega)}, \|z_n - z\|_{L^\infty(\Omega)} \to 0$.

In [2, 3, 4], it is argued that if $\|A_n - A\|_{L^1(\Omega)}, \|z - z_n\|_{L^\infty(\Omega)} \to 0$ and $\{\alpha_n\}_{n=1}^\infty$ is chosen so that $\alpha_n \to 0$ at a rate such that

$$(T_0(A_n u_{\text{exact}}; z_n) - T_0(A_n u_{0,n}; z_n)) \big/ \alpha_n \text{ is bounded,} \tag{3.3}$$

where $u_{0,n}$ is a minimizer of $T_0(A_n u; z_n)$ over $\mathcal{C}$, it follows that

$$\|R_{\alpha_n}(A_n,z_n) - u_{\text{exact}}\|_{L^p(\Omega)} \to 0, \tag{3.4}$$

for $p = 2$ when J is given by (1.4) or (1.5), and for $1 \leq p < d/(d-1)$ when J is given by (1.6). Then R_α satisfies condition 2 of Definition 2.1. However, in [2, 3, 4], the existence of $u_{0,n}$ is not argued. Rather than do this, we can replace (3.3) by

$$\left(T_0(A_n u_{\text{exact}}; z_n) - \inf_{u\in\mathcal{C}} T_0(A_n u; z_n)\right) \Big/ \alpha_n \text{ is bounded,} \tag{3.5}$$

and the arguments of the proofs in [2, 3, 4] go through unchanged. We note that $\inf_{u\in\mathcal{C}} T_0(A_n u; z_n)$ exists due to the fact that $T_0(A_n u; z_n)$ is bounded below. This follows from Jensen's inequality and the properties of the function $x - c\log x$ for $c > 0$:

$$\begin{aligned}T_0(A_n u; z_n) &\geq \|A_n u + \gamma\|_1 - \|z_n\|_\infty \log \|A_n u + \gamma\|_1, \\ &\geq \|z_n\|_\infty - \|z_n\|_\infty \log \|z_n\|_\infty.\end{aligned}$$

For J given by (1.4), however, the arguments of [2] remain incomplete. In particular, from those arguments, we have only that $R_{\alpha_n}(A_n,z_n)$ converges to u_{exact} weakly in $L^2(\Omega)$. But strong convergence can be proved provided (3.5) is replaced by

$$\left(T_0(A_n u_{\text{exact}}; z_n) - \inf_{u\in\mathcal{C}} T_0(A_n u; z_n)\right) \Big/ \alpha_n \to 0. \tag{3.6}$$

This change does not effect the other arguments since (3.6) implies (3.5). Now, to prove strong convergence, we note, as above, that weak convergence in $L^2(\Omega)$, together with $||R_{\alpha_n}(A_n, z_n)||_{L^2(\Omega)} \to ||u_{\text{exact}}||_{L^2(\Omega)}$, implies strong convergence in $L^2(\Omega)$. To prove norm convergence, we note that

$$\inf_{u \in \mathcal{C}} T_0(A_n u, z_n) \leq T_{\alpha_n}(A_n(R_{\alpha_n}(A_n, z_n)), z_n), \tag{3.7}$$

and that by the weak lower semi-continuity of the norm, $||u_{\text{exact}}||^2_{L^2(\Omega)} \leq \liminf ||R_{\alpha_n}(A_n, z_n)||^2_{L^2(\Omega)}$. Thus by (3.7) together with (3.6), we have

$$\begin{aligned}
\liminf ||R_{\alpha_n}(A_n, z_n)||^2_{L^2(\Omega)} &\leq ||u_{\text{exact}}||^2_{L^2(\Omega)} + 2\liminf(\alpha_n^{-1}[T_0(A_n u_{\text{exact}}, z_n) \\
&\qquad - T_0(A_n(R_\alpha(a_n, z_n)), z_n)]) \\
&\leq ||u_{\text{exact}}||^2_{L^2(\Omega)} + 2\liminf(\alpha_n^{-1}[T_0(A_n u_{\text{exact}}, z_n) \\
&\qquad - \inf_{u \in \mathcal{C}} T_0(A_n u, z_n)]) \\
&= ||u_{\text{exact}}||^2_{L^2(\Omega)}.
\end{aligned}$$

Hence we have $||R_{\alpha_n}(A_n, z_n)||^2_{L^2(\Omega)} \to ||u_{\text{exact}}||^2_{L^2(\Omega)}$ and so $R_{\alpha_n}(A_n, z_n)$ converges to u_{exact} strongly in $L^2(\Omega)$.

Thus R_α satisfies condition 2 of Definition 2.1 for J given by (1.4), (1.5), and (1.6).

4. Conclusions

The discussion above constitutes a proof of the following theorem.

Theorem 4.1. *Let $R_\alpha : \mathcal{B} \to \mathcal{C}$ be defined as in (2.2). Then $\{R_\alpha\}_{\alpha>0}$ is a regularization scheme, as defined in Definition 2.1, when J is given by (1.4), (1.5), and (1.6). Moreover, if J is given by (1.4) then* $\text{Range}(R_\alpha) \subset L^2(\Omega)$*; if J is given by (1.5) then* $\text{Range}(R_\alpha) \subset H^1(\Omega)$*; and if J is given by (1.6) then* $\text{Range}(R_\alpha) \subset BV(\Omega)$.

We advocate the use of Definition 2.1 for use in proving analogous results for (2.2) with regularization functionals other than those presented in this paper.

Acknowledgments. The author would like to thank N'djekornom Laobeul and Aaron Luttman for their help with this work. Also, Anatoly Yagola of Moscow State University, Russia, should be mentioned, as it was his course in ill-posed problems at the University of Montana in the fall of 2005 that provided the original impetus for these theoretical explorations.

References

1. R. Acar and C.R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217-1229.

2. J. M. Bardsley and N'djekornom Laobeul, *Tikhonov Regularized Poisson Likelihood Estimation: Theoritical Justification and a Computational Method*, Inverse Problems in Science and Engineering, vol. 16, Issue 2, 2008, pp. 199-215.
3. J. M. Bardsley and N'djekornom Laobeul, *An Analysis of Regularization by Diffusion for Ill-Posed Poisson Likelihood Estimation*, submitted, University of Montana Tech. Report, #23, 2008.
4. J. M. Bardsley and A. Luttman, *Total Variation-Penalized Poisson Likelihood Estimation for Ill-Posed Problems*, accepted in Advances in Computational Mathematics, Special Issue on Mathematical Imaging, University of Montana Tech. Report, #8, 2006.
5. T. Le, R. Chartrand, and T. J. Asaki, *A Variational Approach to Reconstructing Images Corruted by Poisson Noise*, Journal of Mathematical Imaging and Vision, vol. 27, Issue 3, pp. 257-263.
6. Arch W. Naylor and George R. Sell, *Linear Operator Theory in Engineering and Science*, Springer-Verlage 1982.
7. D. L. Snyder, A. M. Hammoud, and R. L. White, *Image recovery from data acquired with a charge-coupled-device camera*, Journal of the Optical Society of America A, **10** (1993), pp. 1014–1023.
8. A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer Academic Publishers, 1990.
9. C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.

Received 12 April, 2007; revised 13 July, 2007

Author information

Johnathan M. Bardsley, Department of Mathematical Sciences, University of Montana, Missoula, MT, 59812., USA.
Email: bardsleyj@mso.umt.edu