



Research article

On the existence and stability of minimizers for generalized Tikhonov functionals with general similarity data

El Mostafa Kalmoun^{1,*} and Fatimah Allami²

¹ Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, P.O. Box 2713, Doha, Qatar

² College of Arts and Sciences in Mahayel Aseer, King Khalid University, Saudi Arabia

* **Correspondence:** Email: ekalmoun@qu.edu.qa; Tel: +97444036594.

Abstract: We study a unified form of Tikhonov regularization for ill-posed problems with a general data similarity term. We discuss sufficient conditions on this generalized Tikhonov functional that guarantee existence and stability of solutions. Furthermore, we show that some particular cases of similarity functionals and regularization techniques can be cast into a unified theoretical framework. In particular, we consider the cases of p -power norm, Bregman distance and mutual information as examples of a data similarity term, and Tikhonov regularization of order one or using a Sobolev norm, total variation penalization and powers of semi-norms associated to closed operators as examples of a regularization term. Finally, we take up the case of the image registration problem.

Keywords: nonlinear ill-posed problems; Tikhonov regularization; existence and stability, image registration

Mathematics Subject Classification: 47J06, 65J20, 49J27

1. Introduction

An inverse problem can be cast into the general form of finding the solution x of a given operator equation

$$F(x) = y, \quad (1.1)$$

where $y \in Y$ is a given data and $F : \Omega \subseteq X \rightarrow Y$ is a general operator, which links two function spaces X and Y that are supposed to be normed spaces.

In recent years, inverse problems have become an area of active research, which is mainly due to the extension of their use to many fields of applications in science and engineering such as mathematical physics, signal processing, medical imaging, computer vision, geophysics, etc. For instance, in medical

imaging, recent applications include computerized tomography, thermoacoustic imaging, electrical impedance tomography; see [18] and references therein.

The distinctive characteristic of these problems is their ill-posedness in the sense of Hadamard [11], which means the problem has either no solution, the solution may not be unique, or the solution does not depend continuously on the data. When the operator F is not surjective, the inverse problem might have no solution, so we instead seek a least residual norm solution. Furthermore, if the solution is not unique (this occurs when F is not injective) we look for a generalized solution called the best approximate to ensure uniqueness. Nevertheless, the main problem of ill-posedness is the instability of this best approximate solution. This can typically be avoided if a stabilization procedure or a regularization method is used.

One of the most popular stabilization procedure is due to Tikhonov [24]. In his method, Tikhonov approximated the solution of the original inverse problem by minimizing both the residual error and the solution norm to penalize large values in the desired solution. Thus the Tikhonov functional related to the inverse problem (1.1) is as follows

$$J_\alpha(x) := \|F(x) - y\|^2 + \alpha\|x\|^2, \quad \forall x \in \Omega, \quad (1.2)$$

where $\alpha > 0$ is a regularization parameter to adjust the weight of the regularization term with respect to the data fitting term. This method is said to be of order zero, while another widely used version replaces the solution norm in the regularization term by the gradient norm. This method is said to be of order one and the energy functional takes the following form

$$J_\alpha(x) := \|F(x) - y\|^2 + \alpha\|\nabla x\|^2, \quad \forall x \in \Omega. \quad (1.3)$$

In his pioneering work, Phillips [19] proposed in the context of integral equations to penalize the large oscillations in the solution by taking the L^2 norm of the second derivative of the solution. This is why the method is also known as Tikhonov-Phillips regularization and the energy functional can be cast into the general form

$$J_\alpha(x) := \|F(x) - y\|^2 + \alpha\|Lx\|^2, \quad \forall x \in \Omega, \quad (1.4)$$

where L is any differential operator of order greater or equal to zero. In some cases, if more regularity in the solution is needed, we can use the Sobolev norm instead of the L^2 norm, which corresponds to a combination of several high order derivatives; see [10]. By the eighties the analysis of Tikhonov regularization for solving linear ill-posed problems was developed [7]. The setting to prove well-posedness (existence, uniqueness and stability) was Hilbert spaces settings for the solution space and the right-hand sides. Recently, the framework has been extended to Banach spaces in connection with Bregman distances; see [1, 3, 12, 16, 21].

On the other hand, the Tikhonov regularization was widely used in many imaging applications. For instance, the first order version (1.3) was applied to the optical flow problem, which gave rise to the well-known method of Horn and Schunck [13]. In many such applications the solution and data spaces were the L^2 space, and the set of admissible solutions was the Sobolev space H^1 . In this case, the method was known to suffer from the over-smoothing effect that can lead to the loss of essential features in the solution. This is mainly due to the use of the L^2 norm as the regularization term in the penalized functional. To overcome this drawback, Rudin, Osher and Fatemi [22] developed the total variation (TV) regularization technique in the context of image denoising and deblurring. The L^2

norm was replaced by the bounded variation semi-norm, which leads to a non-differentiable but still convex regularization term. Despite this theoretical and numerical difficulty in the TV regularization, the method was highly successful in many applications, namely in image processing [4, 5]. In [9], the authors have introduced a method that combines total variation and L^2 regularization to reconstruct piecewise smooth signals. Hofmann *et al* [12] have treated the well-posedness of a generalized Tikhonov functional with a power norm similarity term that includes all the previous methods as special cases

$$J_{\Psi,\alpha}(x) = \|F(x) - y\|^p + \alpha\Psi(x) \quad (1.5)$$

for $1 \leq p < \infty$ and $\Psi : \Omega \subset X \rightarrow [0, \infty]$ is an arbitrary function acting as the regularization procedure. Some recent imaging applications involving non-metric similarity terms has motivated the study of a unified form of the Tikhonov functional with a general similarity data term [8, 20, 23]

$$J_{\Phi,\Psi,\alpha}(x) = \Phi(F(x), y) + \alpha\Psi(x) \quad (1.6)$$

where $\Phi : Y \times Y \rightarrow [0, \infty)$ is a given similarity data functional measuring the error between $F(x)$ and y .

In this paper, we provide weaker sufficient conditions for the existence and stability of the minimizer of the functional $J_{\Phi,\Psi,\alpha}$. While assuming similar conditions on the regularization functional Ψ , our work differs from [17] in taking up a general similarity measure. The obtained results are applied to three known similarity measure; namely, the power norm, the Bregman distance and the mutual information. These functional metrics can be then combined with famous examples of regularization functionals like Tikhonov regularization of order one or using a Sobolev norm, total variation penalization and powers of semi-norms associated to closed operators. We illustrate as well how our results can be applied to the image registration problem, which is an important example of ill-posed inverse problems in computer vision.

The paper consists of four additional sections. In Section 2 we provide some preliminaries. In Section 3 and Section 4, we investigate the existence and stability of minimizers of the generalized Tikhonov functional (1.6). In Section 5 and Section 6, we present some examples of known similarity and regularization functionals. Finally, we consider the case of the image registration problem in Section 7, and conclude our work in Section 8.

2. Preliminaries

We recall the necessary definitions needed in the sequel of this paper. We keep the notations of the previous section to be the same along this paper, and we take X and Y to be two normed spaces unless otherwise stated. We begin by general boundedness and coercivity concepts related to the general functional Ψ instead of the norm function.

Definition 1. A $S \subset \Omega$ is called Ψ -bounded if there exists a constant $c \geq 0$ such that $\Psi(s) \leq c$ for every $s \in S$.

Definition 2. We say that Ω is (weakly) Ψ -complete if each Ψ -bounded sequence of Ω admits a (weakly) convergent subsequence in Ω .

Note that the weak Ψ -completeness holds true for a closed set Ω in a reflexive Banach space X when Ψ is a norm defined on Ω being equivalent or stronger than the norm of X .

Definition 3. We say that a real functional f on Ω is Ψ -coercive if $f(x_n) \rightarrow \infty$ for every sequence $(x_n) \subset \Omega$ with $\Psi(x_n) \rightarrow \infty$. Moreover, a sequence of real functional (f_n) on Ω is Ψ -coercive if $f_n(x_n) \rightarrow \infty$ for every sequence $(x_n) \subset \Omega$ with $\Psi(x_n) \rightarrow \infty$.

The Ψ -coercivity of a real functional f means that $f(x)$ grows infinitely when so does $\Psi(x)$. An example of such a functional f is when it satisfies $f(x) \geq c\Psi(x)$ for $c > 0$. The concept of Ψ -coercivity, which will be used later to show the existence of global minimizers, is closely connected to the Ψ -boundedness of lower level sets.

Proposition 1. If $f : \Omega \rightarrow [-\infty, \infty]$ is Ψ -coercive, then lower level sets of f are Ψ -bounded.

Proof. Suppose $S = \{x \in \Omega : f(x) \leq \alpha\}$ is not Ψ -bounded for some real α . Then there is a sequence $(x_n) \subset S$ satisfying $\Psi(x_n) \rightarrow \infty$. Since $f(x_n) \leq \alpha$, then $f(x_n) \not\rightarrow \infty$. Hence f is not Ψ -coercive. \square

Now, we present a few continuity definitions.

Definition 4. The operator $F : \Omega \rightarrow Y$ is called (weakly) sequentially continuous if for every sequence $(x_n) \in \Omega$ such that $x_n \xrightarrow{(w)} x \in \Omega$, the sequence $F(x_n) \xrightarrow{(w)} F(x)$.

Definition 5. A functional $g : Y \rightarrow [-\infty, \infty]$ is said to be (weakly) lower semicontinuous if for every sequence $(y_n) \in Y$ such that $y_n \xrightarrow{(w)} y \in Y$, we have $g(y) \leq \liminf_{n \rightarrow \infty} g(y_n)$.

Definition 6. A functional $f : \Omega \rightarrow [-\infty, \infty]$ is said to be (weakly) Ψ -subsequentially lower semicontinuous if each Ψ -bounded sequence $(x_n) \subset \Omega$ with $x_n \xrightarrow{(w)} x \in \Omega$ admits a subsequence $(x_{n_j}) \subset (x_n)$ such that $f(x) \leq \liminf_{j \rightarrow \infty} f(x_{n_j})$.

Remark 1. When Ω is (weakly) Ψ -complete, it is clear that every (weakly) lower semicontinuous functional is (weakly) Ψ -subsequentially lower semicontinuous. That the converse is not always true can be seen from the following example. Define

$$f(x) = \begin{cases} x & \text{if } x < 0 \\ x + 1 & \text{if } x \geq 0. \end{cases}$$

It is easy to check that f is not lower semicontinuous. In taking $\Psi(x) = 1/x^2$, we can see that f is continuous on Ψ -bounded set, and hence Ψ -subsequentially lower semicontinuous.

Definition 7. The functional $\Phi : Y \times Y \rightarrow [0, \infty)$ is called Lipschitz continuous on bounded sets of Y if for every bounded set $Z \subset Y$, there exists a constant $c_z \geq 0$ such that

$$|\Phi(y_1, y_2) - \Phi(z_1, z_2)| \leq c_z(\|y_1 - z_1\| + \|y_2 - z_2\|)$$

for every $(y_1, y_2), (z_1, z_2) \in Z \times Z$.

Finally, we recall the definition of uniform convergence of a sequence of functionals.

Definition 8. Let T and (T_n) , $n = 1, 2, \dots$ be functionals defined on Ω . We said that the sequence (T_n) converges uniformly to T on Ψ -bounded sets if for every $c > 0$ and $\varepsilon > 0$, there exists $N = N(c, \varepsilon)$ such that for every $n \geq N$ and every $x \in \Omega$ with $\Psi(x) \leq c$, we have $|T_n(x) - T(x)| < \varepsilon$.

3. Existence theorem

We are now in position to give the main existence theorem of this paper.

Theorem 1. *Suppose that*

- (H1): *F is weakly sequentially continuous,*
- (H2): *Φ is weakly lower semicontinuous in the first argument,*
- (H3): *Ψ is proper and weakly Ψ -subsequentially lower semicontinuous,*
- (H4): *Ω is weakly Ψ -complete.*

Then the functional $J_{\Phi,\Psi,\alpha}$ in (1.6) has a global minimizer.

Proof. Let $(x_n) \subset \Omega$ be a minimizing sequence of $J_{\Phi,\Psi,\alpha}$; that is,

$$\lim_{n \rightarrow \infty} J_{\Phi,\Psi,\alpha}(x_n) = \inf_{x \in \Omega} J_{\Phi,\Psi,\alpha}(x) = J_{\min}. \quad (3.1)$$

First, note that $0 \leq J_{\min} < \infty$. Also, the functional $J_{\Phi,\Psi,\alpha}$ is Ψ -coercive because for every sequence $(z_n) \subset \Omega$ with $\lim_{n \rightarrow \infty} \Psi(z_n) = \infty$, we have

$$J_{\Phi,\Psi,\alpha}(z_n) \geq \alpha \Psi(z_n) \rightarrow \infty.$$

We claim that the sequence (x_n) is Ψ -bounded. Otherwise, there would exist a subsequence $(x_{n_j}) \subset (x_n)$ such that $\Psi(x_{n_j}) \rightarrow \infty$, which yields $J_{\Phi,\Psi,\alpha}(x_{n_j}) \rightarrow \infty$, owing to the Ψ -coercivity of $J_{\Phi,\Psi,\alpha}$, but this contradicts (3.1). Thus, by (H4) there exist a subsequence of (x_n) , denoted the same, and $\bar{x} \in \Omega$ such that $x_n \xrightarrow{w} \bar{x}$. Hence, from (H3), we see that there is another subsequence of (x_n) , denoted again the same, such that

$$\Psi(\bar{x}) \leq \liminf_{n \rightarrow \infty} \Psi(x_n).$$

On the other hand, the combination of (H1) and (H2) tells us that

$$\Phi(F(\bar{x}), y) \leq \liminf_{n \rightarrow \infty} \Phi(F(x_n), y).$$

Therefore

$$\begin{aligned} J_{\Phi,\Psi,\alpha}(\bar{x}) &= \Phi(F(\bar{x}), y) + \alpha \Psi(\bar{x}) \\ &\leq \liminf_{n \rightarrow \infty} \Phi(F(x_n), y) + \alpha \liminf_{n \rightarrow \infty} \Psi(x_n) \\ &\leq \liminf_{n \rightarrow \infty} [\Phi(F(x_n), y) + \alpha \Psi(x_n)] \\ &= \liminf_{n \rightarrow \infty} J_{\Phi,\Psi,\alpha}(x_n) = \lim_{n \rightarrow \infty} J_{\Phi,\Psi,\alpha}(x_n) = J_{\min}, \end{aligned}$$

which means $J_{\Phi,\Psi,\alpha}(\bar{x}) = J_{\min}$, and hence a global minimizer of $J_{\Phi,\Psi,\alpha}$ does exist. \square

Remark 2. *When Ω is Ψ -complete, we can relax the continuity conditions on F , Φ and Ψ to hold true for the strong topology in the related spaces. More precisely, we can replace (H1)-(H4) by the following hypotheses, respectively:*

- (A1): *F is sequentially continuous,*

- (A2): Φ is lower semicontinuous in the first argument,
 (A3): Ψ is proper and Ψ -subsequentially lower semicontinuous,
 (A4): Ω is Ψ -complete.

Remark 3. For a linear operator F , the global minimizer of $J_{\Phi, \Psi, \alpha}$ is unique when $\Phi(\cdot, y)$ and Ψ are convex with at least one of them being strictly convex. For a nonlinear operator, uniqueness can not be guaranteed in the general case.

Remark 4. Theorem 1 recovers Theorem 2.5 in [17] as a special case by taking $\Phi(u, v) = \|u - v\|^2$ and assuming F to be linear.

The previous theorem is a general case for the existence of a minimizer of the Tikhonov functional (1.2).

Corollary 1. Suppose X is a reflexive Banach space and Ω is closed in X . Then the functional (1.2) has a global minimizer provided that F satisfies (H1). Furthermore, when F is linear, the minimizer is unique provided that F is injective or Ω is a Hilbert space.

Proof. For the existence, it suffices to see the functionals $\Phi(z, y) = \|z - y\|^2$ and $\Psi = \|\cdot\|^2$ satisfy assumptions (H2)-(H4) using the reflexivity of X and the weak lower semicontinuity of the norm.

For the uniqueness, when F is injective, the functional $\Phi(\cdot, y)$ is strictly convex, and by convexity of $\Psi(x) = \|x\|^2$, we get the strict convexity of J_α . For the case when (Ω, Ψ) is a Hilbert space, we use the strict convexity of the squared norm $\Psi(x) = \|x\|^2$. \square

4. Stability

In order to study the stability of the generalized Tikhonov functional (1.6), we consider perturbations of the data (F, y) in the inverse problem (1.1) as follows. For $n \in \mathbb{N}$, we let $F_n : \Omega \rightarrow Y$, $y_n \in Y$ and consider the perturbed functionals

$$J_{\Phi, \Psi, \alpha}^{(n)}(x) = \Phi(F_n(x), y_n) + \alpha_n \Psi(x) \quad (4.1)$$

where (α_n) is a positive real sequence converging to α . In general, the minimization of (1.6) might be instable, meaning that minimizers of (4.1) might move away from those of (1.6) even if (F_n, y_n) is enough close to (F, y) . Let's first give a weak stability result for functionals minimization.

Proposition 2. Suppose Ψ satisfies Hypothesis (H4) of Theorem 1. Assume also that J is a weakly Ψ -subsequentially lower semicontinuous real functional on Ω , and (J_n) is a sequence of Ψ -coercive real functionals on Ω which converges uniformly to J on Ψ -bounded sets. Then each minimizer sequence of (J_n) has a weak cluster point that minimizes J .

Proof. Let (x_n) be a sequence of global minimizers of J_n , and x be an arbitrary element of Ω . Since $J_n(x_n) \leq J_n(x)$, it follows that

$$\limsup_{n \rightarrow \infty} J_n(x_n) \leq \limsup_{n \rightarrow \infty} J_n(\bar{x}) = J(\bar{x}) < \infty.$$

On the basis of the Ψ -coercivity of J_n , it's clear that $\Psi(x_n)$ is bounded above; hence it is bounded, which means the sequence (x_n) is Ψ -bounded. In combining this with the hypothesis (H4), we see that

there exist $x^* \in \Omega$ and a subsequence of (x_n) , denoted the same way, such that $x_n \xrightarrow{w} x^*$. Since (x_n) is Ψ -bounded and J is weakly Ψ -subsequentially lower semicontinuous, there is a subsequence, denoted again (x_n) , such that $J(x^*) \leq \liminf_{n \rightarrow \infty} J(x_n)$. Recalling that (J_n) converges uniformly to J on Ψ -bounded sets, we get $J(x_n) - J_n(x_n) \rightarrow 0$. Consequently,

$$\begin{aligned} J(x^*) &\leq \liminf_{n \rightarrow \infty} J(x_n) \leq \limsup_{n \rightarrow \infty} J(x_n) \\ &\leq \limsup_{n \rightarrow \infty} (J(x_n) - J_n(x_n)) + \limsup_{n \rightarrow \infty} J_n(x_n) \\ &= \limsup_{n \rightarrow \infty} J_n(x_n) \\ &\leq J(\bar{x}), \end{aligned}$$

which holds for every $x \in \Omega$. We conclude that x^* is a global minimizer of J . \square

As a particular case, we get a weak stability result for the minimizer of the generalized Tikhonov functional (1.6).

Theorem 2. Suppose (F_n) satisfies (H1) such that (F_n) converges uniformly to F on Ψ -bounded sets, and let $y_n \rightarrow y$. In addition to the hypotheses of Theorem 1, assume also that Φ is Lipschitz continuous on Ψ -bounded sets. Then $\operatorname{argmin} J_{\Phi, \Psi, \alpha}^{(n)} \neq \emptyset$, $\operatorname{argmin} J_{\Phi, \Psi, \alpha} \neq \emptyset$ and each sequence in the former set has a weak cluster point in the latter one.

Proof. On the basis of Theorem 1, it's clear to see that $\operatorname{argmin} J_{\Phi, \Psi, \alpha}^{(n)} \neq \emptyset$ and $\operatorname{argmin} J_{\Phi, \Psi, \alpha} \neq \emptyset$. It is enough to check that the functionals $J_{\Phi, \Psi, \alpha}$ and $J_{\Phi, \Psi, \alpha}^{(n)}$ satisfy the hypotheses of Proposition 2.

First, in applying similar arguments as in the proof of Theorem 1, we can easily show that $J_{\Phi, \Psi, \alpha}$ is weakly Ψ -subsequentially lower semicontinuous. For the Ψ -coercivity of $(J_{\Phi, \Psi, \alpha}^{(n)})$, we have for any sequence (x_n) in Ω satisfying $\Psi(x_n) \rightarrow \infty$,

$$J_{\Phi, \Psi, \alpha}^{(n)}(x_n) \geq \alpha_n \Psi(x_n) \rightarrow \infty.$$

Finally, to show the uniform convergence of $(J_{\Phi, \Psi, \alpha}^{(n)})$ to $J_{\Phi, \Psi, \alpha}$ on Ψ -bounded sets, we let M be a Ψ -bounded set in Ω . Owing to the Lipschitz continuity of Φ on Ψ -bounded sets, we see that for every $x \in M$

$$\begin{aligned} |J_{\Phi, \Psi, \alpha}^{(n)}(x) - J_{\Phi, \Psi, \alpha}(x)| &= |\Phi(F_n(x), y_n) + \alpha_n \Psi(x) - \Phi(F(x), y) - \alpha \Psi(x)| \\ &\leq |\Phi(F_n(x), y_n) - \Phi(F(x), y)| + |(\alpha_n - \alpha)| \Psi(x) \\ &\leq c(\|F_n(x) - F(x)\| + \|y_n - y\|) + |(\alpha_n - \alpha)| \Psi(x) \end{aligned}$$

where $c > 0$ is a Lipschitz constant. This yields $(J_{\Phi, \Psi, \alpha}^{(n)})$ converges uniformly to $J_{\Phi, \Psi, \alpha}$ on M because so does (F_n) to F , and taking into account $y_n \rightarrow y$ and $\alpha_n \rightarrow \alpha$. \square

Remark 5. We observe that the weak stability result of Theorem 2 is satisfied if

1. The space X is a reflexive Banach space and Ω is closed in X .
2. F is weakly sequentially continuous.
3. The functionals Φ and Ψ are norms.

Taking into account Remark 2, the following strong stability result may be proved in much the same way as Theorem 2.

Theorem 3. *Suppose (F_n) satisfies (A1) such that (F_n) converges uniformly to F on Ψ -bounded sets, and let $y_n \rightarrow y$. In addition to the hypotheses of Remark 2, assume also that Φ is Lipschitz continuous on Ψ -bounded sets. Then $\operatorname{argmin} J_{\Phi, \Psi, \alpha}^{(n)} \neq \emptyset$, $\operatorname{argmin} J_{\Phi, \Psi, \alpha} \neq \emptyset$ and each sequence in the former set has a cluster point in the latter one.*

5. Examples of similarity measures

In this section, we present two examples of similarity measures that correspond to particular choices of the functional Φ in (1.6).

5.1. Norms

The most known example of a similarity measure is the power norm function

$$\Phi(z, y) = \|z - y\|^p, \quad 1 \leq p < \infty. \quad (5.1)$$

Letting $p = 2$, we get the quadratic data term in the standard Tikhonov functional (1.2). The functional Φ satisfies Hypothesis (H2) of Theorem 1 by the weak lower semicontinuity of the norm.

For the weak stability result in Theorem 2, we need to show that Φ is Lipschitz continuous on bounded sets assuming Ψ to be a norm on X . Let $Z \subset Y$ be a bounded set. First, we have $\|z\| \leq \|y\| + \|z - y\|$. Using the p -duality map, we can show that

$$\|z\|^p \leq \|y\|^p + p\|z - y\|\|z\|^{p-1}.$$

Then, by exchanging the role of z and y , we get

$$|\|z\|^p - \|y\|^p| \leq pC^p\|z - y\|$$

where C is the boundedness constant of Z . Hence,

$$\begin{aligned} |\Phi(z_1, y_1) - \Phi(z_2, y_2)| &\leq pC^p\|z_1 - y_1 - z_2 + y_2\| \\ &\leq pC^p(\|z_1 - z_2\| + \|y_1 - y_2\|) \end{aligned}$$

for all z_1, y_1, z_2, y_2 in Z .

5.2. Bregman distance

Bregman distance [2] was introduced in 1967 in the context of convex programming. Here we recall a generalized definition of this pseudo-distance related to convex functionals on a Banach space. Suppose Y is a Banach space and let $f : Y \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function. Let also v^* be a subgradient of f at $v \in Y$; that is an element of the subdifferential

$$\partial f(v) = \{w \in Y^* \mid f(v) + \langle w, u - v \rangle \leq f(u) \quad \forall u \in Y\}.$$

Since the function J is convex, finite and continuous at one point of U , then $\Omega_{\mathcal{B}}(f) = \{v : \partial f(v) \neq \emptyset\}$ the Bregman domain of f satisfies $\Omega_{\mathcal{B}}(f) = Y$; see [6] (Proposition 5.2 in Chapter I). The Bregman distance is then defined for $u, v \in Y$ by

$$B_f^{v^*}(u, v) := f(u) - f(v) - \langle v^*, u - v \rangle.$$

The symmetric version of the Bregman distance is given by

$$B_f^s(u, v) := B_f^{u^*}(v, u) + B_f^{v^*}(u, v) = \langle u^* - v^*, u - v \rangle.$$

We take $\Phi(u, v) = B_f^{v^*}(u, v)$. By the definition of subgradients, the nonnegativity of Φ is clear. Moreover, the weak lower semicontinuity of Φ in the first argument results from the convexity and continuity of f .

For stability with respect to perturbation of right hand data y only, we need to assume a Lipschitz condition on f . Owing to the local boundedness of the subdifferential sets of f , the similarity functional Φ will be Lipschitz with respect to the first argument and this will be sufficient for the proof of Theorem 2. In case of perturbations in both the operator F and the data y , it is useful to use the symmetric version of the Bregman distance to get the joint Lipschitz condition of Φ .

6. Examples of regularization terms

In this section we give different known regularization terms that meet the conditions we have imposed on the functional Ψ in (1.6).

6.1. Tikhonov regularization of order one

Instead of penalizing large components in the solution by choosing the penalizer $\Psi(x) = \|x\|^2$ in the ordinary Tikhonov functional (1.2), we can penalize large deviations in the solution by penalizing the squared norm of the gradient.

Let $U \subset \mathbb{R}^d$ be open and bounded with a smooth boundary, $X = L^2(U)$ and $\Omega = H^1(U)$. We take $\Psi(x) = \|\nabla x\|_{L^2(U)}^2$ in the functional (1.6). As the L^2 norm is weakly lower semicontinuous, the functional Ψ satisfies Hypothesis (H3). Note also that Ψ satisfies also Hypothesis (H4) in view of the Poincaré inequality and the reflexivity of the Sobolev space Ω .

Under the hypotheses (H1) and (H2), if the operator F is linear and the similarity measure Φ is convex in the first argument, then the minimizer is unique because Ψ is already strictly convex as the composition of a strictly convex function and a linear operator.

6.2. Tikhonov regularization using Sobolev norm

In case of seeking further regularity in the solution, we can apply a higher-order regularization by making use of the Sobolev norm. In particular, another different form of Tikhonov regularization arises by choosing the penalizer as $\Psi(x) = \|x\|_{W^{k,p}}$ where $W^{k,p}$ is a Sobolev space [15].

Let $U \subset \mathbb{R}^d$ be open and bounded with a smooth boundary, $X = L^p(U)$ and $\Omega = W^{k,p}(U)$ with $1 < p < \infty$. To prove that Ψ satisfies Hypotheses (H3) and (H4), let us consider $(x_n) \subset W^{k,p}(U)$ to be a Ψ -bounded sequence such that $x_n \xrightarrow{w} x$ in $W^{k,p}(U)$ and let $\alpha \in \mathbb{Z}^n$ be a multi-index such that $|\alpha| \leq k$. It

follows that $D^\alpha x_n \xrightarrow{w} D^\alpha x$ in $L^p(U)$. By the weak lower semicontinuity of the L^p norm, we have

$$\|D^\alpha x\|_{L^p(U)} \leq \liminf_{n \rightarrow \infty} \|D^\alpha x_n\|_{L^p(U)}.$$

Thus

$$\begin{aligned} \Psi(x) &= \|x\|_{W^{k,p}(U)} = \left(\sum_{|\alpha| \leq k} \|D^\alpha x\|_{L^p(U)}^p \right)^{1/p} \\ &\leq \liminf_{n \rightarrow \infty} \left(\sum_{|\alpha| \leq k} \|D^\alpha x_n\|_{L^p(U)}^p \right)^{1/p} \\ &= \liminf_{n \rightarrow \infty} \|x_n\|_{W^{k,p}(U)} = \liminf_{n \rightarrow \infty} \Psi(x_n). \end{aligned}$$

This shows (H3). Moreover, since $W^{k,p}(U)$ is compactly embedded in $L^p(U)$, then there exist a subsequence $(x_{n_j}) \subset (x_n)$ and $x \in W^{k,p}(U)$ such that $x_{n_j} \xrightarrow{w-L^p} x$. Hence (H4) holds true.

Once again, the uniqueness is guaranteed in case of the convexity of $\Phi(., y)$ and linearity of F provided that Hypotheses (H1) and (H2) are satisfied because Ψ is strictly convex.

6.3. Total variation penalization

The second popular regularization technique after Tikhonov method is total variation penalization. The method was first introduced by Rudin *et al* [22] in the context of image denoising, and since then it achieved a highly success in many imaging applications [4, 14]. We will follow the same arguments employed in [17] to show that this penalization term satisfies the conditions of Theorem 1 and Theorem 2.

Let $U \subset \mathbb{R}^d$ be convex and bounded with Lipschitz continuous boundary such that $d \geq 2$ and $X = L^p(U)$ with $1 \leq p \leq \frac{d}{d-1}$. We consider Ω to be the space of functions of bounded variation (BV) on U :

$$\Omega = BV(U) = \{x \in L^1(U) : \varphi_0(x) < \infty\}$$

with

$$\varphi_0(x) = \sup_{y \in V} \int_U (-x \operatorname{div} y) ds$$

and

$$V = \{y \in C_0^1(U; \mathbb{R}^d) : |y(s)| \leq 1, \forall s \in U\}.$$

Note that if $x \in C^1(U)$, then $\varphi_0(x) = \int_U |\nabla x| ds$. The BV norm is then defined for $x \in \Omega$ by $\|x\|_{BV(U)} = \|x\|_{L^1(\Omega)} + \varphi_0(x)$, and we take $\Psi(x) = \|x\|_{BV(U)}$. Let $(x_n) \subset \Omega$ be a Ψ -bounded sequence satisfying $x_n \xrightarrow{w-L^p} x \in \Omega$; hence $u_n \xrightarrow{w-L^1} u$ (since $p \geq 1$). Since the norm $\|\cdot\|_{L^1(U)}$ is weakly lower semicontinuous, we get

$$\|x\|_{L^1(\Omega)} \leq \liminf_{n \rightarrow \infty} \|x_n\|_{L^1(\Omega)}. \quad (6.1)$$

On the other hand, from $x_n \xrightarrow{w-L^p} x$, we can see that

$$\int_U (-x \operatorname{div} y) ds = \lim_{n \rightarrow \infty} \int_U (-x_n \operatorname{div} y) ds$$

$$\leq \liminf_{n \rightarrow \infty} \int_U (-x_n \operatorname{div} y) ds.$$

By taking supremum on both sides, we obtain

$$\varphi_0(x) \leq \liminf_{n \rightarrow \infty} \varphi_0(x_n).$$

The combination of the last inequality with (6.1) tells us that

$$\begin{aligned} \Psi(x) &= \|x\|_{BV(U)} = \|x\|_{L^1(U)} + \varphi_0(x) \\ &\leq \liminf_{n \rightarrow \infty} \|x_n\|_{L^1(U)} + \liminf_{n \rightarrow \infty} \varphi_0(x_n) \\ &\leq \liminf_{n \rightarrow \infty} (\|x_n\|_{L^1(U)} + \varphi_0(x_n)) \\ &= \liminf_{n \rightarrow \infty} \|x_n\|_{BV(U)} = \liminf_{n \rightarrow \infty} \Psi(x_n). \end{aligned}$$

This proves (H3). Furthermore, let $(x_n) \subset \Omega$ be a Ψ -bounded sequence. As $BV(U)$ is compactly embedded in $L^p(U)$ for $1 \leq p \leq \frac{d}{d-1}$, there is a subsequence (x_{n_j}) of (x_n) such that $x_{n_j} \xrightarrow{w-L^p} x$, which confirms (H4).

If the operator F is linear and $\Phi(\cdot, y)$ is strictly convex, then the minimizer is unique; otherwise we cannot guarantee uniqueness because the $\|\cdot\|_{BV}$ norm is not strictly convex.

6.4. Regularization with q -powers of seminorms

We consider here a very general form of Tikhonov regularization [17], which corresponds to the choice of $\Psi(x) = \|Lx\|^q$ in the generalized Tikhonov functional (1.6) where $L : \Omega \subset X \rightarrow Y$ is a closed linear operator with a weakly closed range $R(L)$. Note that a closed and bounded below linear operator on a Banach space has a closed range. Here we suppose that X is a Hilbert space and Y is a reflexive Banach space. Let us denote the null space of L by $N(L)$, the orthogonal space of a subspace M in X by M^\perp , the orthogonal projection of X onto M by P_M , and finally the Moore-Penrose generalized inverse of L by L^\dagger .

Let $(x_n) \subset \Omega$ be a Ψ -bounded sequence such that $x_n \xrightarrow{w} x \in \Omega$. It is easily seen that the sequence (Lx_n) is bounded in the weakly closed set $R(L)$. Therefore it has a weakly convergent subsequence (Lx_{n_j}) to an element z of $R(L)$ because Y is reflexive. Now, by the weak continuity of L^\dagger , we have $P_{N(L)^\perp}(x_{n_j}) = L^\dagger(Lx_{n_j}) \xrightarrow{w} L^\dagger z$. Thus $P_{N(L)}(x_{n_j}) = x_{n_j} - P_{N(L)^\perp}x_{n_j} \xrightarrow{w} x - L^\dagger z$. Since $N(L)$ is closed, we get $x - L^\dagger z \in N(L)$; and therefore

$$0 = L(x - L^\dagger z) = Lx - LL^\dagger z = Lx - P_{\overline{R(L)}}(z) = Lx - z.$$

Hence $Lx = z$. It follows that

$$\Psi(x) = \|Lx\|^q = \|z\|^q \leq \liminf_{j \rightarrow \infty} \|Lx_{n_j}\|^q = \liminf_{j \rightarrow \infty} \Psi(x_{n_j}).$$

Therefore Ψ satisfies (H3). Furthermore, since Ψ -bounded sequences in Ω are bounded in X which is reflexive, then Hypothesis (H4) holds true as well.

For a linear operator F and a convex functional Φ , the minimum is unique by use of the strict convexity of Ψ as the q -th power ($q > 1$) of a strictly convex norm.

Remark 6. The hypotheses on L hold for most differential operators on Sobolev spaces.

7. Application to image registration

In this section we take up the image registration problem, which is subject to intense research efforts due to its important applications in remote sensing, computer vision, medical imaging, etc. Given a reference image $R : U \rightarrow [0, 1]$ and a template image $T : U \rightarrow \mathbb{R}$ of the same object, defined on a bounded domain $U \in \mathbb{R}^d$, the problem consists in finding a spatial displacement field $x : U \rightarrow U$ such that $T(s + x(s)) = R(s)$ for all s in U . This is a well-known example of ill-posed inverse problems, which can be cast into the form of a functional equation as

$$T(x) = R.$$

Under a variational framework, this problem is solved by minimizing the generalized Tikhonov functional (1.6) for $F = T$, $y = R$, Ω and Y are given spaces of displacement fields and images, respectively.

The choice of the similarity measure Φ can be based either on the alignment of the geometrical structures or the gray levels of the two images. An extraordinary successful similarity metric in the context of multi-modal registration is the mutual information proposed by Viola and Wells [25]. In assuming $\xi_1^x := T(x)$ and $\xi_2 := R$ are random variables in Y with respective probability densities p_T and p_R , the mutual information measures the Kullback-Leibler distance between the joint distribution $p_{T,R}(\xi_1^x, \xi_2)$ and the product distribution $p_T(\xi_1^x)p_R(\xi_2)$

$$\mathcal{D}(T(x), R) = KL(p_{T,R}(\xi_1^x, \xi_2), p_T(\xi_1^x)p_R(\xi_2))$$

where $KL(p_1, p_2) = \int p_1(s) \log \frac{p_1(s)}{p_2(s)} ds$ for any densities p_1 and p_2 . This functional is known to be well-defined and positive. Also, if

$$Y = \{y \in L^\infty(U) : \|y\|_\infty \leq A\}$$

for a large positive constant A , then $\mathcal{D}(T(x), R) \leq C$ where C is a constant that depends on A . It can be shown (see [23]) that $\Phi(u, v) = -\mathcal{D}(u, v) + C$ is Lipschitz continuous on $Y \times Y$ in any L^q norm for $q \geq 1$, which means that Assumption (A2) is satisfied as well. On the other hand, the operator $F = T$ is sequentially continuous when the space of displacement fields Ω is continuously embedded into $L^q(U, \mathbb{R}^d)$. A frequent case of such a space would be a Banach space that is compactly embedded into $L^q(U, \mathbb{R}^d)$. In combining the mutual information metric with any choice of a regularization functional from the previous section, we get the existence and stability of the image registration problem.

8. Conclusions

We have considered a unified form of Tikhonov regularization for ill-posed problems which involve a general similarity data term. Under weaker sufficient conditions, we have proven the existence and stability of global minima of the generalized Tikhonov functional. In some special cases, we have precised the additional conditions to ensure the uniqueness of the global minimum. We have also illustrated how known examples of similarity measures and regularization terms meet the existence and stability conditions we have imposed on the generalized Tikhonov functional. Finally, we have briefly demonstrated that our results can be applied to prove the existence and stability of solutions to the image registration problem.

Acknowledgment

The authors would like to thank two anonymous reviewers for their valuable suggestions and corrections.

References

1. S. W. Anzengruber, *The Discrepancy Principle for Tikhonov Regularization in Banach spaces: Regularization Properties and Rates of Convergence*. Südwestdeutscher Verlag für Hochschulschriften, 2012.
2. L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.*, **7** (1976), 200–217.
3. M. Burger, S. Osher, Convergence rates of convex variational regularization, *Inverse Probl.*, **20** (2004), 1411.
4. V. Caselles, A. Chambolle, M. Novaga, Total variation in imaging, *Handb. Math. Methods Imaging*, **2015** (2015), 1455–1499.
5. T. Chan, S. Esedoglu, F. Park, A. Yip, Recent developments in total variation image restoration, *Math. Models Comput. Vision*, **17** (2005).
6. I. Ekeland, R. Temam, *Convex analysis and variational problems*, **28** (1999).
7. H. W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Springer Science Business Media, **375** (1996).
8. J. Flemming, Theory and examples of variational regularization with non-metric fitting functionals, *J. Inverse Ill-Posed Probl.*, **18** (2010), 677–699.
9. A. Gholami, S. M. Hosseini, A balanced combination of Tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals, *Signal Processing*, **93** (2013), 1945–1960.
10. C. Groetsch, The theory of Tikhonov regularization for Fredholm equations. 104p, *Boston Pitman Publication*, 1984.
11. J. Hadamard, Lectures on Cauchy's problem in linear partial differential equations, Yale Univ. Press. New Haven, 1923.
12. B. Hofmann, B. Kaltenbacher, C. Poeschl, O. Scherzer, A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators, *Inverse Problems*, **23** (2007), 987.
13. B. K. Horn, B. G. Schunck, Determining optical flow, *Artif. Intel.*, **17** (1981), 185–203.
14. E. M. Kalmoun, An investigation of smooth tv-like regularization in the context of the optical flow problem, *J. Imaging*, **4** (2018), 31.
15. P. Kazemi, R. J. Renka, Tikhonov regularization using Sobolev metrics, *Electron. J. Differ. Eq.*, 2014.
16. S. Kindermann, Convex Tikhonov regularization in Banach spaces: New results on convergence rates, *J. Inverse Ill-posed Probl.*, **24** (2016), 341–350.

17. G. Mazziere, R. Spies, K. Temperini, Existence, uniqueness and stability of minimizers of generalized Tikhonov–Phillips functionals, *J. Math. Anal. Appl.*, **396** (2012), 396–411.
18. F. D. M. Neto, A. J. da Silva Neto, *An Introduction to Inverse Problems with Applications*, Springer Science & Business Media, 2012.
19. D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. ACM (JACM)*, **9** (1962), 84–97.
20. C. Pöschl, *Tikhonov Regularization with General Residual Term*, PhD thesis, University of Innsbruck, Austria, 2008.
21. E. Resmerita, Regularization of ill-posed problems in Banach spaces: Convergence rates, *Inverse Probl.*, **21** (2005), 1303.
22. L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenom.*, **60** (1992), 259–268.
23. S. Salzo, *Variational Regularization for Image Registration: Theory and Algorithms*, PhD thesis, Dipartimento di Informatica e Scienze dell’Informazione, 2012.
24. A. N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, *Soviet Math.*, **4** (1963), 1035–1038.
25. P. Viola, W. M. Wells III, Alignment by maximization of mutual information, *Int. J. Comput. Vision*, **24** (1997), 137–154.



AIMS Press

© 2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)