УДК 519.67

O. Kravchuk, G. Kriukova

DOI: 10.18523/2617-70805202257-61

REGULARIZATION BY DENOISING FOR INVERSE PROBLEMS IN IMAGING

In this work, a generalized scheme of regularization of inverse problems is considered, where a priori knowledge about the smoothness of the solution is given by means of some self-adjoint operator in the solution space. The formulation of the problem is considered, namely, in addition to the main inverse problem, an additional problem is defined, in which the solution is the right-hand side of the equation. Thus, for the regularization of the main inverse problem, an additional inverse problem is used, which brings information about the smoothness of the solution to the initial problem. This formulation of the problem makes it possible to use operators of high complexity for regularization of inverse problems, which is an urgent need in modern machine learning problems, in particular, in image processing problems. The paper examines the approximation error of the solution of the initial problem using an additional problem.

Keywords: inverse problems.

Introduction

Solving modern machine learning tasks requires development of new methods of solving corresponding inverse problems. Majority of real-world inverse problems are ill-posed and therefore require regularization. For some digital signal processing tasks, such as image de-noising, image restoration, super-resolution, image improvement, the choice of regularization technique is non-trivial, whereas significantly influences the corresponding solution.

In our work we study generalized regularization scheme for inversion of image transforms. For inverse problem

$$Ax = y$$

we consider Bayesian approach, or maximum a posteriori probability (MAP) estimate, which finds such an x, that maximises the conditional probability $p(x|y)$. According to Bayes rule

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx} \propto p(y|x)p(x),$$

therefore maximisation of $p(x|y)$ corresponds to the following problem:

$$\arg\min_x(-\log p(y|x) - \log p(x)).$$

Obviously, real probability distribution functions are unknown. Therefore instead of it we solve the following heuristics

$$\hat{x} = \arg\min_x\{l(x,y) + \alpha\rho(x)\}, \qquad (1)$$

where $l(x,y)$ is a loss function and $\rho(x)$ is a regularization term.

Let's slightly modify (1):

$$\hat{x} = \arg\min_{x,v}\{l(x,y) + \alpha\rho(v)\}, x = v.$$

It allows us to apply Alternating Direction Method of Multipliers (ADMM) from the paper [2], using Lagrangian:

$$L_\lambda(x,v,u) = l(x,y)+\alpha\rho(v)+\frac{\lambda}{2}\|x-v+u\|^2-\frac{\lambda}{2}\|u\|^2.$$

It leads to iterative solving following minimization tasks till convergence:

$$\hat{x} \longleftarrow \arg\min_x L(x,\hat{v},u)$$
$$\hat{v} \longleftarrow \arg\min_x L(\hat{x},v,u)$$
$$u \longleftarrow u + (\hat{x} - \hat{v})$$

or after redefining variables in terms of (1)

$$\hat{x} = \min_x l(x,y) + \beta\|x-v\|^2,$$
$$\hat{v} = \min_v \alpha\rho(v) + \beta\|x-v\|^2.$$

In such a way, instead of one inverse problem with regularization scheme we've got two interconnected minimization problems, iterative solving of which allows us to find solution for the initial problem. Having some initial x_0 and v_0 we iterate

$$x_{i+1} = \min_x l(x,y) + \beta\|x-v_i\|^2,$$
$$v_{i+1} = \min_v \alpha\rho(v) + \beta\|x_i-v\|^2.$$

Let's consider some operator $D : X \mapsto X$, that preserves x as a solution, i.e.

$$AD(x,\sigma) = y,$$

for example, for super-resolution task instead of D a de-noising operator may be used.

This allows to see all the setting from another perspective: we have inverse problem $Ax = y$ and its corresponding loss-function $l(x, y)$, and some other problem with loss-function $\rho(x)$.

In such a way, instead of one inverse problem we get two interconnected problems, where second one brings prior information to the first one. As an additional problem, any prior information may be used. For example, some external classifier for image generation improvement problem or denoising problem as a regularization for super-resolution problem, etc.

The idea to use image denoisers as a mechanism behind the regularization term underlies the Regularization by Denoising (RED) framework [5]

$$\rho_{RED}(x) \triangleq \frac{1}{2}\langle x, x - f(x)\rangle,$$

where $f(\cdot)$ is the denoiser of choice.

This idea has been broadened to the following setting, called Regularization by Denoising via Fixed-Point Projection (RED-PRO) [1].

$$\hat{x}_{RED-PRO} = \arg\min_x l(x, y),$$
$$s.t. \quad \|x - f(x)\|^2 = 0.$$

It can be seen as a solving of inverse problem with additional inverse problem $x = f(x)$ as a regularization.

The same framework may be seen as a regularization by means of regularization term $\rho(x)$ is $\alpha\rho(x) = \alpha x^T[x - D(x, \sigma)]$. Under mild conditions (differentiability, local homogeneity, and symmetric Jacobian for D) gradient descent may be applied to get the solution:

$$x_{k+1} = x_k - \mu[A^T(Ax_k - y) - \alpha[x_k - D(x_k, \sigma)]].$$

In [1] it has been shown, that Plug-and-Play Prior (PnP) proximal gradient method considered in [4] is a special case of Regularization by Denoising via Fixed-Point Projection (RED-PRO), the convergence of both frameworks to globally optimal solutions has been proven as a result of the convergence analysis and the study of the solutions of both PnP and RED frameworks [1].

Another classical approach to solving inverse problem $Ax = b$ is the method of Tikhonov-Phillips regularization in Hilbert scales, where a regularized approximation x_α^δ is defined as the solution of the minimization problem

$$\min_{x \in \mathcal{D}(B^s)} \|Ax - y^\delta\|^2 + \alpha\|B^s x\|^2,$$

where $\alpha > 0$ is the regularization parameter, $B : \mathcal{D}(B) \subset X \to X$ is an unbounded densely defined self-adjoint strictly positive definite operator and s is some non-negative real number to be chosen properly to influence the properties of the regularized approximation x_α^δ.

In [7] it was shown that under the assumptions

$$\|B^p \hat{x}\| \leq E$$

and

$$m\|B^{-a}x\| \leq \|Ax\| \leq M\|B^{-a}x\|$$

with some constants E, m and M, the Tikhonov-Phillips regularized approximation x_α^δ of problem $Ax = y$ provides order optimal error bounds

$$\|x_\alpha^\delta - \hat{x}\| = O(\delta^{p/(a+p)})$$

for $s \geq (p - a)/2$, in the case that α is chosen ***a priori*** by $\alpha = c\delta^{2(a+s)/(a+p)}$ with some constant $c > 0$.

In the paper we study the approach to solving inverse problem with regularization by means of additional inverse problem with fixed-point projection, that may be seen as a smoothing condition [6].

General Regularization Scheme

In this paper we consider ill-posed problem

$$Ax = y, \tag{2}$$

where $A : X \to Y$ is a bounded linear operator between real Hilbert spaces X and Y with non-closed range $\mathcal{R}(A)$. Let's denote the inner producy by $\langle \cdot, \cdot \rangle$ and the corresponding norm on the Hilbert spaces by $\| \cdot \|$.

We assume, that the operator A is injective and that y belongs to $\mathcal{R}(A)$. It implies that (2) has a unique solution $\hat{x} \in X$. Suppose that instead of exact data y we have an available data $y^\delta \in Y$ such that

$$\|y - y^\delta\| \leq \delta \tag{3}$$

for some known noise level δ. Since $\mathcal{R}(A)$ is assumed to be non-closed, the solution $\hat{x}$ does not depend continuously on the exact y and available data y^δ. Hence, the problem (2) is ill-posed and therefore requires the regularization. Regularization is reconstruction of the solution of problem with inexact data using additional information, for example, (A1) subjective information concerning the smoothness of $\hat{x}$ and (A2) objective information concerning the smoothing property of the operator A.

To formulate the smoothing properties we use densely defined unbounded self-adjoint strictly positive operator $B : X \to X$ and some index function φ.

Definition 1. Function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is called *index function* if it is continuous and strictly increasing with $\varphi(0+) = 0$.

Assumption A1: For some $p > 0$ and $E < \infty$, the solution $\hat{x}$ of the problem (2) satisfies $\|B^p(x)\| \leq E$.

Assumption A2: There exists some index function φ with properties:

(i) there exists a constant $m > 0$ with $m\|\sqrt{\varphi(B^{-2})}x\| \leq \|Ax\|$ for all $x \in X$

(ii) for p as in Assumption A1, the function $\lambda\varphi(\lambda^{1/p}) : (0, \|B^{-2p}] \to \mathbb{R}^+$ is convex

Definition 2. *General regularization scheme* in Hilbert space is defined as

$$x_\alpha = B^{-s} g_\alpha(T^*T)T^*y,$$

$$x_\alpha^\delta = B^{-s} g_\alpha(T^*T)T^*y^\delta$$

with $T = AB^{-s}$ for some $s \geq 0$ and piece-wise continuous $g_\alpha : (0, \|T\|^2] \to \mathbb{R}$ with the property

$$\lim_{\alpha \to 0+} g_\alpha(\lambda) = 1/\lambda.$$

For further analysis we make additional assumption about the function g_α.

Assumption A3: There exist positive constants γ and β such that

$$\begin{aligned} \sup_{\lambda>0} \quad \sqrt{\lambda}|g_\alpha(\lambda)| &\leq \gamma/\sqrt{\alpha}, \\ \sup_{\lambda>0} \quad \lambda|g_\alpha(\lambda)| &\leq 1, \\ \sup_{\lambda>0} \quad \sqrt{\lambda}|1-\lambda g_\alpha(\lambda)| &\leq \beta\sqrt{\alpha}, \\ \sup_{\lambda>0} \quad \lambda|1-\lambda g_\alpha(\lambda)| &\leq 1. \end{aligned}$$

Different regularization methods are characterized by corresponding functions g_α.

For example, ordinary Tikhonov-Phillips regularization in Hilbert space is defined by $g_\alpha = 1/(\lambda+\alpha)$. In this case Assumption A3 is satisfied with $\gamma = 1/2$ and $\beta = 1/2$. For Tikhonov-Phillips regularization of order m in Hilbert space the function g_α is defined as following:

$$g_\alpha = \frac{1}{\lambda}\left(1 - \left(\frac{\alpha}{\lambda+\alpha}\right)^m\right)$$

with $\gamma = \sqrt{m}$ and $\beta = 1$.

Spectral method of regularization in Hilbert space is defined by $g_\alpha(\lambda) = \frac{1}{\max\{\lambda,\alpha\}}$ with $\gamma = 1$ and $\beta = 2/\sqrt{27}$. Asymptotical regularization in Hilbert space is defined by $g_\alpha(\lambda) = \frac{1}{\lambda}(1 - e^{-\lambda/\alpha})$, $\gamma = 1$, $\beta = 1/\sqrt{2e}$. Finally, iterative regularization in Hilbert space, also known as Landweber iteration, are defined by

$$g_\alpha(\lambda) = \frac{1}{\lambda}\left(1 - (1-\lambda)^{1/\alpha}\right),$$

for $\gamma = 1$ and $\beta = 1/\sqrt{2e}$.

In [6] it was shown that under Assumption A2 the regularized approximation x_α^δ with $s = p$ is order optimal if α is chosen *a priori*.

Theorem 1 ([6])**.** *Let x_α^δ be regularized approximation defined by general regularization scheme (see Definition 2) with s chosen by $s = p$ and let assumptions A1 and A3 be satisfied. Then, for $\alpha = \frac{\delta^2}{E^2}$,*

$$\|x_\alpha^\delta - \hat{x}\| \leq$$

$$(\gamma+1)\sup_{x\in X}\{\|x\| : \|B^r x\| \leq E, \|Ax\| \leq c\delta\}$$

with $c = \frac{\beta+1}{\gamma+1}$. If, in addition, assumption A2 is satisfied, then

$$\|x_\alpha^\delta - \hat{x}\| \leq (\gamma+1)E\sqrt{\psi_p^{-1}\left(\frac{c^2\delta^2}{m^2E^2}\right)},$$

where $\psi_p(\cdot)$ is defined as $\psi_p(\lambda) = \lambda\phi(\lambda^{1/p})$.

It implies Mair's convergence rate result for the method of Tikhonov-Phillips regularization. In fact, the second error bound of Theorem 1 shows the order optimality of the regularized approximation x_α^δ (see [6]).

Another inverse problem as a regularisation

Let's come back to the initial inverse problem

$$Ax = y.$$

And let's consider another inverse problem

$$Dq = x.$$

Then general regularization scheme defines the following solution:

$$q_\alpha^\delta = g_\alpha(D^*D)D^*x.$$

In [3] within the proof of Proposition 2.8 it was shown that for the whole class of regularization families the estimate of regularization error has the form

$$\|\hat{q} - q_\alpha^\delta\| \leq R\hat{\gamma}\varphi(\alpha) + \gamma_{-1/2}\frac{\delta}{\sqrt{\alpha}}.$$

The first term here depends on the smoothness of the solution, and in the statistical spirit we agreed to call it the bias. Then the second term is the variance, and its order $\delta/\sqrt{\alpha}$ is the same for all regularization families under consideration.

Then we have

$$Dq_\alpha^\delta = Dg_\alpha(D^*D)D^*x \approx x,$$

thus let's denote the following operator:

$$B = Dg_\alpha(D^*D)D^*x : X \to X.$$

Easy to see, that B is self-adjoint strictly positive definite operator. Then all the theoretical results from the previous section are applicable to this problem setting.

Conclusions

Regularization of ill-posed operator equations in Hilbert scales is usually studied under the assumption that the operator A involved in the equation and the operator B generating the Hilbert scale are related by some operator-valued index function ϕ. In the classical paper [7] of Natterer, such a relation that characterizes the smoothing properties of A relative to the operator B^{-1} has been expressed in terms of power functions (see [6]). Extensions to general index functions have been considered in Mair's paper [8] for the case of high-order regularization in Tikhonov-Phillips regularization method. In our paper we compare classical results for a general regularization scheme to the case of regularization by means of denoiser operator. Another accomplishment of this paper is the justification of error bounds in the light of general index functions ϕ. It is important to note that the general regularization scheme requires neither any knowledge of the index function ϕ nor any knowledge of the solution smoothness measured against the Hilbert scale. Nevertheless, it automatically provides an order optimal solution for the considered ill-posed problem.

References

1. C. Regev, E. Michael and M. Peyman, "Regularization by Denoising via Fixed-Point Projection (RED-PRO)", SIAM Journal on Imaging Sciences. Society for Industrial and Applied Mathematics. **14** (3) (2021), https://doi.org/10.1137/20M1337168
2. S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers", Found. Trends Mach. Learn. **3** (1), 1–122 (2011).
3. S. Pereverzyev, *Selected Topics of the Regularization Theory* (Springer International Publishing. Cham., 2014).
4. V. Singanallur Venkatakrishnan, Charles A. Bouman and Brendt Wohlberg, "Plug-and-play priors for model based reconstruction", in: *2013 IEEE Global Conference on Signal and Information Processing* (2013), pp. 945–948.
5. Yaniv Romano, Michael Elad and Peyman Milanfar, "The little engine that could: Regularization by denoising (RED)", SIAM Journal on Imaging Sciences. **10** (4), 1804–1844 (2017).
6. M. T. Nair, S. V. Pereverzyev and U. Tautenhahn, "Regularization in Hilbert scales under general smoothing conditions", Inverse Problems. IOP Publishing Ltd. **21** (6), 1851 (2005).
7. F. Natterer, "Error bounds for Tikhonov regularization in Hilbert scales", Appl. Anal. **18**, 29–37 (1984).
8. B. A. Mair, "Tikhonov regularization for finitely and infinitely smoothing operators", SIAM J. Math. Anal. **25**, 135–147 (1994).

Кравчук О. М., Крюкова Г. В.

РЕГУЛЯРИЗАЦІЯ ЗА ДОПОМОГОЮ ВИДАЛЕННЯ ШУМУ В ОБЕРНЕНИХ ЗАДАЧАХ ОБРОБКИ ЗОБРАЖЕНЬ

У цій роботі розглянуто узагальнену схему регуляризації обернених задач, де апріорне знання про гладкість розв'язку дано за допомогою деякого самоспряженого оператора в просторі розв'язків. Розглянуто постановку задачі, коли окрім основної оберненої задачі визначено додаткову задачу, в якій шуканий розв'язок є правою частиною рівняння. Таким чином, для регуляризації основної оберненої задачі використовується додаткова обернена задача, яка привносить до початкової задачі інформацію про гладкість розв'язку. Така постановка задачі дає можливість використовувати оператори високої складності для регуляризації обернених задач, що є нагальною потребою в сучасних задачах машинного навчання, зокрема, в задачах обробки зображень. В роботі досліджено похибку апроксимації розв'язку початкової задачі за допомогою додаткової задачі.

Ключові слова: обернені задачі.

Матеріал надійшов 27.10.2022