PRIN2022_MORIGI dal titolo Inverse Problems in the Imaging Sciences (IPIS) 2022ANC8HL - CUP J53D23003670006

Title: Learned regularization for Inverse Problems

The name *inverse problem* is widely used to denote the task of reconstructing an unknown quantity from indirect, possibly corrupted and incomplete measurements. The mathematical model which describes the measurement generation is supposed to be known, and is characterized by a *forward operator*, the (linear or nonlinear) map that connects the unknown quantity with the observable one, and a noise model. Despite the forward operator is usually well-known and grounded on solid physical and mathematical bases, its inversion might be challenging: namely, the solution *x* associated with a measurement *y*, if it exists, may not be unique, and may notably change even in the presence of small perturbations of y. For these reasons, such problems are referred to as *ill-posed* problems. Several examples of inverse problems appear in many real-world situations, ranging from signal processing, biomedical imaging, geoscience and astronomy. From a mathematical standpoint, the forward operator *A* usually involves, for instance, linear integral transforms (such as Fourier transform or Radon transform, which are the backbones of Magnetic Resonance Imaging and X-rays Computed Tomography, respectively), or nonlinear maps between the coefficients and the solutions of partial differential equations (such as in the Calderon problem, involved in Electrical Impedance Tomography).

A common strategy to tackle ill-posedness is by means of regularization theory, which studies the approximation of inverse problems by means of families of well-posed problems. One prominent class of methods, among the many strategies proposed in the last decades, is the so-called *variational regularization*, which approximate the original inverse problem with a family of minimization problems, where the functional to be minimized is a combination between a data fidelity term (measuring the distance between *A(x)* and *y*) and a regularization term, penalizing undesired features in the solution. Variational regularization schemes, together with the many other traditional alternatives in literature, are referred to as *model-driven* methods, because they essentially rely on the knowledge of the forward operator, and because they try to encode prior knowledge on *x* through interpretable mathematical models (e.g. smoothness, sparsity, etc.).

On the other side, the possibility to acquire large collections of data is bringing the attention of the inverse problems community towards *data-driven* methods, that are able to capture bespoke structures in data without requiring the introduction of refined and complicated models. Such techniques can be framed in the context of statistical learning, whose task is to identify a function connecting an input variable to an output one, taking advantage of many samples of training data. For this reason, these methods are referred to as learning-based, and model-blind, since they can spare the knowledge of the forward model. The growing interest for learning methods in diverse fields of inverse problems, e.g. image denoising, super-resolution, limited-angle tomography and edge detection is motivated by the great success and high performance they can achieve. In this context, neural networks (and especially deep ones) are an essential tool and represent a class of models where learning algorithms are successfully set, thanks to strong numerical evidence about their expressivity and versatility.

Despite the wide and successful range of application of data-driven methods, numerous aspects call for a deeper investigation. For example, the performances of neural networks in many applications significantly depend on the choice of the network architecture, which is often arbitrary, or defined by means of trial-and-error procedures. Moving from black-box models to networks that are (fully or partially) theoretically justified, guaranteeing comparable numerical results, is a field of growing interest. Further aspects of theoretical investigation are the dependence of the learned method from the data set used to generate it, and its stability with respect to perturbations of the input data. These and many related issues are the main focuses of the so-called *Mathematics of Machine Learning*, a lively field of mathematics at the intersection of numerical analysis, functional analysis, optimization and statistics. A rigorous mathematical investigation of learning techniques is essential, especially if applied to inverse problems, in which the intrinsic ill-posedness of the model might cause instable reconstructions, hallucinations, or poor generalization properties.

The current proposal focuses on hybrid learning techniques for inverse problems, particularly in the framework of *learned regularization*. This name denotes a growing family of techniques which are inspired by classical regularization methods, while taking advantage of tools from statistical learning and neural networks. For example, the *Plug and Play* paradigm is inspired by variational regularization, and particularly by iterative algorithms to solve the associated minimization problems, and replaces the model-based proximal functional appearing in the iterations by neural network specifically trained on a denoising task. Other prominent examples include *unrolling*, *generative models*, and *network Tikhonov*.

The main advantage of those technique is their ability to bridge the gap between variational and learned techniques, allowing both for theoretical justifications and competitive numerical performances. The goal is to develop and motivate novel learned regularization schemes, proving mathematical guarantees on them, and applying them on problems closely related with real-world applications. We put particular interest on the following topics:

- **Networks tailored to inverse problems applications:**
  In many inverse problems of interest, the input and output variables might have a rather complicate nature. For example, in problems related with partial differential equations such as EIT, the coefficients to be identified are discretized on computational meshes, which require a graph representation and a more sophisticated set of tools than the ones usually employed in imaging. In the same application, the measured data consists of a linear operator, whose efficient representation and manipulation is an open issue. We are interested in drafting neural architecture which are tailored to deal with such inputs and outputs. Moreover, whenever the original problem is formulated in a continuous framework, discretization invariance (or even an infinite-dimensional setting) of such algorithm is a desired property.
- **Theoretical deduction, or motivation, of the architecture:**
  We want to propose neural architecture of interpretable nature, which are *aware* of the underlying physical model and incorporate it in the learning process. We are willing to trade a slight drop in accuracy, compared with state-of-the-art black box models, in order to develop techniques which are theoretically motivated, and for which it is possible to prove mathematical properties, such as convergence of the algorithm and stability of the outputs.
- **Learning techniques as regularization strategies:**
  One theoretical property we are looking for, when developing novel hybrid techniques, is their ability to regularize the original inverse problem, namely, to mitigate its ill-posedness. A common way to do so is to prove that, if the noise level on the measurement is reduced, the regularized solution converges to the true solution of the inverse problem. At the moment, there exist few theoretical studies in this direction, which is nevertheless a rather important task in other fields of artificial intelligence, under the general name of “Trustworthy AI”.
- **Generalization properties of the trained networks:**
  Another crucial property we are looking for is the ability of the learned regularizers to *generalize*, namely, to extract useful information from the training dataset and perform well also when applied to unseen data. The mathematical study of such a property nowadays relies on powerful statistical tools, but still has some vulnerabilities when applied to high-dimensional and ill-posed problems. This is also a topic of growing interest also outside the presented field, and it falls in the category of “Robust AI”.