Acta Numerica (2019), pp. 1–174
doi:10.1017/S0962492919000059

Solving inverse problems using data-driven models

Simon Arridge
Department of Computer Science,
University College London,
Gower Street, London WC1E 6BT, UK
E-mail: S.Arridge@cs.ucl.ac.uk

Peter Maass
Department of Mathematics,
University of Bremen, Postfach 330 440,
28344 Bremen, Germany
E-mail: pmaass@math.uni-bremen.de

Ozan Öktem
Department of Mathematics,
KTH – Royal Institute of Technology,
SE-100 44 Stockholm, Sweden
E-mail: ozan@kth.se

Carola-Bibiane Schönlieb
Department of Applied Mathematics and Theoretical Physics,
Cambridge University, Wilberforce Road,
Cambridge, CB3 0WA, UK
E-mail: C.B.Schoenlieb@damtp.cam.ac.uk

Recent research in inverse problems seeks to develop a mathematically coherent foundation for combining data-driven models, and in particular those based on deep learning, with domain-specific knowledge contained in physical–analytical models. The focus is on solving ill-posed inverse problems that are at the core of many challenging applications in the natural sciences, medicine and life sciences, as well as in engineering and industrial applications. This survey paper aims to give an account of some of the main contributions in data-driven inverse problems.

CONTENTS

1. Introduction

In several areas of science and industry there is a need to reliably recover a hidden multi-dimensional model parameter from noisy indirect observations. A typical example is when imaging/sensing technologies are used in medicine, engineering, astronomy and geophysics. These so-called inverse problems are often ill-posed, meaning that small errors in data may lead to large errors in the model parameter, or there are several possible model parameter values that are consistent with observations. Addressing ill-posedness is critical in applications where decision making is based on the recovered model parameter, for example in image-guided medical diagnostics. Furthermore, many highly relevant inverse problems are large-scale: they involve large amounts of data and the model parameter is high-dimensional.

Traditionally, an inverse problem is formalized as solving an equation of the form

$$g = \mathcal{A}(f) + e.$$

Here $g \in Y$ is the measured data, assumed to be given, and $f \in X$ is the model parameter we aim to reconstruct. In many applications, both g and f are elements in appropriate function spaces Y and X, respectively. The mapping $\mathcal{A}\colon X \to Y$ is the forward operator, which describes how the model parameter gives rise to data in the absence of noise and measurement errors, and $e \in Y$ is the observational noise that constitutes random corruptions in the data g. The above view constitutes a knowledge-driven approach, where the forward operator and the probability distribution of the observational noise are derived from first principles.

Classical research on inverse problems has focused on establishing conditions which guarantee that solutions to such ill-posed problems exist and on methods for approximating solutions in a stable way in the presence of noise (Engl, Hanke and Neubauer 2000, Benning and Burger 2018, Louis 1989, Kirsch 2011). Despite being very successful, such a knowledge-driven approach is also associated with some shortcomings. First, the forward model is always an approximate description of reality, and extending it might be challenging due to a limited understanding of the underlying physical or technical setting. It may also be limited due to computational complexity. Accurate analytical models, such as those based on systems of non-linear partial differential equations (PDEs), may reach a numerical complexity beyond any feasible real-time potential in the foreseeable future. Second, most applications will have inputs which do not cover the full model parameter space, but stem from an unknown subset or obey an unknown stochastic distribution. The latter shortcoming in particular has led to the advance of methods that incorporate information about the structure of the parameters to be determined in terms of sparsity assumptions (Daubechies, Defrise and De Mol 2004, Jin and Maass 2012*b*) or stochastic models (Kaipio and Somersalo 2007, Mueller and Siltanen 2012). While representing a significant advancement in the field of inverse problems, these models are, however, limited by their inability to capture very bespoke structures in data that vary in different applications.

At the same time, data-driven approaches as they appear in machine learning offer several methods for amending such analytical models and for tackling these shortcomings. In particular, deep learning (LeCun, Bengio and Hinton 2015), which has had a transformative impact on a wide range of tasks related to artificial intelligence, ranging from computer vision and speech recognition to playing games (Igami 2017), is starting to show its impact on inverse problems. A key feature in these methods is the use of *generic* models that are adapted to specific problems through learning against example data (training data). Furthermore, a common trait in the success stories for deep learning is the abundance of training data and the explicit agnosticism from *a priori* knowledge of how such data are generated. However, in many scientific applications, the solution method needs to be robust and there is insufficient training data to support an entirely data-driven approach. This seriously limits the use of entirely data-driven approaches for solving problems in the natural and engineering sciences, in particular for inverse problems.

A recent line of development in computational sciences combines the seemingly incompatible data- and knowledge-driven modelling paradigms. In the context of inverse problems, ideally one uses explicit knowledge-driven models when there are such available, and learns models from example data using data-driven methods only when this is necessary. Recently several algorithms have been proposed for this combination of model- and data-driven approaches for solving ill-posed inverse problems. These results are still primarily experimental and lack a thorough theoretical foundation; nevertheless, some mathematical concepts for treating data-driven approaches for inverse problems are emerging.

This survey attempts to provide an overview of methods for integrating data-driven concepts into the field of inverse problems. Particular emphasis is placed on techniques based on deep neural networks, and our aim is to pave the way for future research towards providing a solid mathematical theory. Some aspects of this development are covered in recent reviews of inverse problems and deep learning, for instance those of McCann, Jin and Unser (2017), Lucas, Iliadis, Molina and Katsaggelos (2018) and McCann and Unser (2019).

1.1. Overview

This survey investigates algorithms for combining model- and data-driven approaches for solving inverse problems. To do so, we start by reviewing some of the main ideas of knowledge-driven approaches to inverse problems, namely functional analytic inversion (Section 2) and Bayesian inversion (Section 3), respectively. These knowledge-driven inversion techniques are derived from first principles of knowledge we have about the data, the model parameter and their relationship to each other.

Knowledge- and data-driven approaches can now be combined in several different ways depending on the type of reconstruction one seeks to compute and the type of training data. Sections 4 and 5 represent the core of the survey and discuss a range of inverse problem approaches that introduce data-driven aspects in inverse problem solutions. Here, Section 4 is the data-driven sister section to functional analytic approaches in Section 2. These approaches are primarily designed to combine data-driven methods with functional analytic inversion. This is done either to make functional analytic approaches more data-driven by appropriate parametrization of these approaches and adapting these parametrizations to data, or to accelerate an otherwise costly functional analytic reconstruction method.

Many reconstruction methods, however, are not naturally formulated within the functional analytic view of inversion. An example is the posterior mean reconstruction, whose formulation requires adopting the Bayesian view of inversion. Section 5 is the data-driven companion to Bayesian inversion in Section 3, and surveys methods that combine data- and knowledge-driven methods in Bayesian inversion. The simplest is to apply data-driven post-processing of a reconstruction obtained via a knowledge-driven method. A more sophisticated approach is to use a learned iterative scheme that integrates a knowledge-driven model for how data are generated into a data-driven method for reconstruction. The latter is done by unrolling a knowledge-driven iterative scheme, and both approaches, which compute

Cambridge University Press & Assessment
978-1-108-47868-7 — Acta Numerica 2019
Arieh Iserles
Excerpt
More Information

statistical estimators, can be combined with forward operators that are partially learned via a data-driven method.

The above approaches come with different trade-offs concerning demands on training data, statistical accuracy and robustness, functional complexity, stability and interpretability. They also impact the choice of machine learning methods and algorithms for training. Certain recent – and somewhat anecdotal – topics of data-driven inverse problems are discussed in Section 6, and exemplar practical inverse problems and their data-driven solutions are presented in Section 7.

Within data-driven approaches, deep neural networks will be a focus of this survey. For an introduction to deep neural networks the reader might find it helpful to consult some introductory literature on the topic. We recommend Courville, Goodfellow and Bengio (2017) and Higham and Higham (2018) for a general introduction to deep learning; see also Vidal, Bruna, Giryes and Soatto (2017) for a survey of work that aims to provide a mathematical justification for several properties of deep networks. Finally, the reader may also consult Ye, Han and Cha (2018), who give a nice survey of various types of deep neural network architectures.

Detailed structure of the paper. In Section 2 we discuss functional analytic inversion methods, and in particular the mathematical notion of ill-posedness (Section 2.3) and regularization (Section 2.4) as a means to counteract the latter. A special focus is on variational regularization methods (Sections 2.5–2.7), as those reappear in bilevel learning in Section 4.3 in the context of data-driven methods for inverse problems.

Statistical – and in particular Bayesian – approaches to inverse problems are described in Section 3. In contrast to functional analytic approaches (Section 2.4), in Bayesian inversion (Section 3.1) the model parameter is a random variable that follows a prior distribution. A key difference between Bayesian and functional analytic inversion is that in Bayesian inversion an approximation to the whole distribution of the model parameter conditioned on the measured data (posterior distribution) is computed, rather than a single model parameter as in functional analytic inversion. This means that reconstructed model parameters can be derived via different estimates of its posterior distribution (a concept that we will encounter again in Section 5, and in particular Section 5.1.2, where data-driven reconstructions are phrased as results of different Bayes estimators), but also that uncertainty of reconstructed model parameters can be quantified (Section 3.2.5). When evaluating different reconstructions of the model parameter – which is again important when defining learning, *i.e.* optimization criteria for inverse problem solutions – aspects of statistical decision theory can be used (Section 3.3). Also, the parallel concept of regularization, introduced in Section 3 for the functional analytic approach, is outlined in Section 3.2 for statistical approaches. The difficult problem of selecting a prior distribution for the model parameter is discussed in Section 3.4.

In Section 4 we present some central examples of machine learning combined with functional analytic inversion. These encompass classical parameter choice rules for inverse problems (Section 4.1) and bilevel learning (Section 4.3) for parameter learning in variational regularization methods. Moreover, dictionary learning is discussed in Section 4.4 as a companion to sparse reconstruction methods in Section 2.7, but with a data-driven dictionary. Also, the concept of a black-box denoiser, and its application to inverse problems by decoupling the regularization from the inversion of the data, is presented in Section 4.6. Two recent approaches that use deep neural network parametrizations for data-driven regularization in variational inversion models are investigated in Section 4.7. In Section 4.9 we discuss a range of learned optimization methods that use data-driven approximations as a means to speed up numerical computation. Finally, in Section 4.10 we introduce a new idea of using the recently introduced concept of deep inverse priors for solving inverse problems.

In Section 5 learning data-driven inversion models are phrased in the context of statistical regularization. Section 5.1.2 connects back to the difficulty in Bayesian inversion of choosing an appropriate prior (Section 3.4), and outlines how model learning can be used to compute various Bayes estimators. Here, in particular, fully learned inversion methods (Section 5.1.3), where the whole inversion model is data-driven, are put in context with learned iterative schemes (Section 5.1.4), in which data-driven components are interwoven with inverse model assumptions. In this context also we discuss post-processing methods in Section 5.1.5, where learned regularization together with simple knowledge-driven inversion methods are used sequentially. Section 5.2 addresses the computational bottleneck of Bayesian inversion methods by using learning, and shows how one can use learning to efficiently sample from the posterior.

Section 6 covers special topics of learning in inverse problems, and in Section 6.1 includes task-based reconstruction approaches that use ideas from learned iterative reconstruction (Section 5.1.4) and deep neural networks for segmentation and classification to solve joint reconstruction-segmentation problems, learning physics-based models via neural networks (Section 6.2.1), and learning corrections to forward operators by optimization methods that perform joint reconstruction-operator correction (Section 6.2).

Finally, Section 7 illustrates some of the data-driven inversion methods discussed in the paper by applying them to practical inverse problems. These include an introductory example on inversion of ill-conditioned linear systems to highlight the intricacy of using deep learning for inverse problems as a black-box approach (Section 7.1), bilevel optimization from Section 4.3 for parameter learning in TV-type regularized problems and variational models with mixed-noise data fidelity terms (Section 7.2), the application of learned iterative reconstruction from Section 5.1.4 to computed tomography (CT) and photoacoustic tomography (PAT) (Section 7.3), adversarial regularizers from Section 4.7 for CT reconstruction as an example of variational regularization with a trained neural network as a regularizer (Section 7.4), and the application of deep inverse priors from Section 4.10 to magnetic particle imaging (MPI) (Section 7.5).

In Section 8 we finish our discussion with a few concluding remarks and comments on future research directions.

2. Functional analytic regularization

Functional analysis has had a strong impact on the development of inverse problems. One of the first publications that can be attributed to the field of inverse problems is that of Radon (1917). This paper derived an explicit inversion formula for the so-called Radon transform, which was later identified as a key component in the mathematical model for X-ray CT. The derivation of the inversion formula, and its analysis concerning missing stability, makes use of operator formulations that are remarkably close to the functional analysis formulations that would be developed three decades later.

2.1. The inverse problem

There is no formal mathematical definition of an inverse problem, but from an applied viewpoint such problems are concerned with determining causes from desired or observed effects. It is common to formalize this as solving an operator equation.

Definition 2.1. An *inverse problem* is the task of recovering the model parameter $f_{\text{true}} \in X$ from measured data $g \in Y$, where

$$g = \mathcal{A}(f_{\text{true}}) + e. \tag{2.1}$$

Here, X (model parameter space) and Y (data space) are vector spaces with appropriate topologies and whose elements represent possible model parameters and data, respectively. Moreover, $\mathcal{A}\colon X \to Y$ (forward operator) is a known continuous operator that maps a model parameter to data in absence of observation noise and $e \in Y$ is a sample of a Y-valued random variable modelling the observation noise.

In most imaging applications, such as CT image reconstruction, elements in X are images represented by functions defined on a fixed domain $\Omega \subset \mathbb{R}^d$ and elements in Y represent imaging data by functions defined on a fixed manifold $\mathbb{M}$ that is given by the acquisition geometry associated with the measurements.

2.2. Introduction to some example problems

In the following, we briefly introduce some of the key inverse problems we consider later in this survey. All are from imaging, and we make a key distinction between (i) *image restoration* and (ii) *image reconstruction*. In the former, the data are a corrupted (*e.g.* noisy or blurry) realization of the model parameter (image) so the reconstruction and data spaces coincide, whereas in the latter the reconstruction space is the space of images but the data space has a definition that is problem-dependent. As we will see when discussing data-driven approaches to inverse problems in Sections 4 and 5, this differentiation is particularly crucial as the difference between image and data space poses additional challenges to the design of machine learning methods. Next, we describe very briefly some of the most common operators that we will refer to below. Here the inverse problems in Sections 2.2.1–2.2.3 are image restoration problems, while those in Sections 2.2.4 and 2.2.5 are examples of image reconstruction problems.

2.2.1. Image denoising

The observed data are the ideal solution corrupted by additive noise, so the forward operator in (2.1) is the identity transform $\mathcal{A} = \mathrm{id}$, and we get

$$g = f_{\text{true}} + e, \tag{2.2}$$

In the simplest case the distribution of the observational noise is known. Furthermore, this distribution may in more advanced problems be correlated, spatially varying and of mixed type.

In Section 7.2 we will discuss bilevel learning of total variation (TV)-type variational models for denoising of data corrupted with mixed noise distributions.

2.2.2. Image deblurring

The observed data are given by convolution with a known filter function K together with additive noise, so (2.1) becomes

$$g = f_{\text{true}} * K + e. \tag{2.3}$$

Any inverse problem of the type (2.1) with a linear forward operator that is translation-invariant will be of this form.

In the absence of noise, the inverse problem (*image deconvolution*) is exactly solvable by division in the Fourier domain, *i.e.* $f_{\text{true}} = \mathcal{F}^{-1}[\mathcal{F}[g]/\mathcal{F}[K]]$, provided that $\mathcal{F}[K]$ has infinite support in the Fourier domain. In the presence of noise, the estimated solution is corrupted by noise whose frequency spectrum is the reciprocal of the spectrum of the filter K. The distribution of the observational also has the same considerations as in (2.2). Finally, extensions include the case of a spatially varying kernel and the case where K is unknown (*blind deconvolution*).

2.2.3. Image in-painting

Here, the observed data represents a noisy observation of the true model parameter $f_{\text{true}}\colon \Omega \to \mathbb{R}$ restricted to a fixed measurable set $\Omega_0 \subset \mathbb{R}^n$:

$$g = f_{\text{true}}\big|_{\Omega_0} + e. \tag{2.4}$$

In the above, $f_{\text{true}}\big|_{\Omega_0}$ is the restriction of f_{true} to Ω_0. Solutions take different forms depending on the size of connected components in Ω_0. Extensions include the case where Ω_0 is unknown or only partially known.

2.2.4. Computed tomography (CT)

The simplest physical model for CT assumes mono-energetic X-rays and disregards scattering phenomena. The model parameter is then a real-valued function $f\colon \Omega \to \mathbb{R}$ defined on a fixed domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ for two-dimensional CT and $d = 3$ for three-dimensional CT) that has unit mass per volume. The forward operator is the one given by the Beer–Lambert law:

$$\mathcal{A}(f)(\omega, x) = \mathrm{e}^{-\mu \int_{-\infty}^{\infty} f(x+s\omega)\,\mathrm{d}s}. \tag{2.5}$$

Here, the unit vector $\omega \in S^{d-1}$ and $x \in \omega^{\perp}$ represent the line $\ell : s \mapsto x + s\omega$ along which the X-rays travel, and we also assume f decays fast enough for the integral to exist. In medical imaging, μ is usually set to a value that approximately corresponds to water at the X-ray energies used. The above represents pre-logarithm (or pre-log) data, and by taking the logarithm (or log) of data, one can recast the inverse problem in CT imaging to one where the forward model is the linear ray transform:

$$\mathcal{A}(f)(\omega, x) = \int_{-\infty}^{\infty} f(x + s\omega)\,\mathrm{d}s. \tag{2.6}$$

For low-dose imaging, pre-log data are Poisson-distributed with mean $\mathcal{A}(f_{\text{true}})$, where $\mathcal{A}$ is given as in (2.5), that is, $g \in Y$ is a sample of $\mathbb{g} \sim \text{Poisson}(\mathcal{A}(f_{\text{true}}))$. Thus, to get rid of the non-linear exponential in (2.5), it is common to take the log of data. With such post-log data the forward operator is linear and given as in (2.6). A complication with such post-log data is that the noise model becomes non-trivial, since one takes the log of a Poisson-distributed random variable (Fu *et al.* 2017). A common approximate noise model for post-log data is (2.1), with observational noise e which is a sample of a Gaussian or Laplace-distributed random variable.

In the case of *complete* data, that is, where a full angular set of data is measured, an exact inverse is obtained by the (Fourier-transformed) data backprojected on the same lines as used for the measurements and scaled by the absolute value of the spatial frequency, followed by the inverse Fourier transform. Thus, as in deblurring, the noise is amplified, but only linearly in spatial frequency, making the problem mildly ill-posed. Extensions include the *emission tomography* problem (single photon emission computed tomography (SPECT) and positron emission tomography (PET)) where the line integrals are exponentially attenuated by a function μ that may be unknown. A major challenge in tomography is to consider *incomplete* data, in particular the case where only a subset of lines is measured. This problem is much more ill-posed.

See Sections 7.3, 7.4 and 7.6 for instances of CT reconstruction that use deep neural networks in the solution of the inverse problem.

2.2.5. Magnetic resonance imaging (MRI)

The observed data are often considered to be samples of the Fourier transform of the ideal signal, so the MRI image reconstruction problem is an inverse problem of the type (2.1), where the forward operator is given as a discrete sampling operator concatenated with the Fourier transform. A correct description of the problem takes account of the complex-valued nature of the data, which implies that when e is normally distributed then the noise model of $|\mathcal{F}^{-1}[g]|$ is Rician. As in CT, the case of under-sampled data is of high practical importance. In MRI, the subsampling operator has to consist of connected trajectories in Fourier space but is not restricted to straight lines.

Extensions include the case of *parallel MRI* where the forward operator is combined with (several) spatial sensitivity functions. More exact forward operators take account of other non-linear physical effects and can reconstruct several functions in the solution space.

2.3. Notion of ill-posedness

A difficulty in solving (2.1) is that the solution is sensitive to variations in data, which is referred to as ill-posedness. More precisely, the notion of ill-posedness is usually attributed to Hadamard, who postulated that a well-posed problem must have three defining properties, namely that (i) it has a solution (existence) that is (ii) unique and that (iii) depends continuously on the data g (stability). Problems that do not fulfil these criteria are ill-posed and, according to Hadamard, should be modelled differently (Hadamard 1902, Hadamard 1923).

For example, instability arises when the forward operator $\mathcal{A}\colon X \to Y$ in (2.1) has an unbounded or discontinuous inverse. Hence, every non-degenerate compact operator between infinite-dimensional Hilbert spaces whose range is infinite naturally leads to ill-posed inverse problems. Slightly more generally, one can prove that continuous operators with non-closed range yield unbounded inverses and hence lead to ill-posed inverse problems. This class includes non-degenerate compact operators as well as convolution operators on unbounded domains.