MAT-62006 Inverse Problems, Part 2: Statistical Inversion

S. Pursiainen

Department of Mathematics, Tampere University of Technology

Spring 2015

TAMPERE UNIVERSITY OF TECHNOLOGY

Overview

The statistical inversion approach is based on the following principles:

- The inverse problem is presented as a statistical quest for information.
- We have directly observable quantities and others that cannot be observed.
- In inverse problems, some of the unobservable quantities are of primary interest.
- These quantities depend on each other through forward and inverse models.

Overview

The statistical inversion approach is based on the following principles:

1. All variables included in the model are modelled as random variables.
2. The randomness describes our degree of information concerning their realizations.
3. The degree of information concerning these values is coded in the probability distributions.
4. The solution of the inverse problem is the posterior probability distribution (Bayesian inverse model).

Overview

1. Forming the posterior density often is not enough to enlighten possible realizations of the unknown, since the shape of the posterior distribution is often likely to be somewhat tricky.
2. Factors that can make exploration of the posterior difficult include, among other things, a complicated forward model, a high number dimensions, existence of several local maxima, peaked structure, and/or weak prior information of the unknown.
3. Due to these difficulties numerical optimization and sampling methods are typically needed to find a reconstruction of the unkown based on the posterior.

Bayes' Formula

- Assume that we are measuring a quantity $y \in \mathbb{R}^m$ in order to get information about another quantity $x \in \mathbb{R}^n$.
- In order to relate these two quantities, we need a data prediction (forward) model of their dependence, e.g. of the form
$$y = f(x, \varepsilon),$$
where $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ is the model function and $\varepsilon \in \mathbb{R}^k$ is a noise vector.
- In statistical inverse problems, all parameters are viewed as random variables denoted here by upper case letters (e.g. X and Y).
- Realizations of the random variables are denoted by lower case letters (e.g. x and y).
- Thus, the model $y = f(x, \varepsilon)$ would lead to a relation
$$Y = f(X, E).$$

Bayes' Formula

- Before performing the measurement of Y there often is some prior (*a priori*) information, i.e. a foresight, of the unknown variable X.
- Bayesian approach relies on the assumption that the prior information can be coded into a prior probability density $x \to \pi_{pr}(x)$.
- Joint probability density $\pi(x, y)$ can be written in the following two ways: $\pi(x, y) = \pi_{pr}(x)\pi(y \mid x)$ and $\pi(x, y) = \pi(y)\pi(x \mid y)$.
- These together imply the classical Bayes' formula

$$\pi(x \mid y) = \frac{\pi_{pr}(x)\pi(y \mid x)}{\pi(y)}.$$

Bayes' Formula

- In Bayes' formula, $\pi(x \mid y)$ is known as the posterior probability density of x given the measurement y, which is in this context a solution of an inverse problem.
- The likelihood function of measuring y with a certain realization of x is given by $\pi(y \mid x)$, and strongly related to the forward model.
- Density $\pi(y)$ can be treated as a constant, when y is given, and therefore, Bayes' formula is often expressed in the form
$$\pi(x \mid y) \propto \pi_{pr}(x)\pi(y \mid x).$$
- It is possible (e.g. due to nullspace effects) that the likelihood function cannot be scaled so that $\int \pi(y \mid x)\, dy = 1$, i.e. it is not always a well-defined probability density.

Bayes' Formula

Example 1 (Matlab)

Assume that the goal is to detect an electric charge located at the point $x = (r, \varphi) = (0.5, 0.5)$ (polar coordinates) inside an unit disk D using n equally spaced sensors lying on the boundary of D with one of the sensors placed on positive x-axis. Sensor i experiences the electric potential $y_i = 1/d_i + n_i$, in which d_i is the distance to the charge and n_i is a zero mean Gaussian noise random variable with standard deviation σ. From $n_i = y_i - 1/d_i$ it follows that $\pi(n_i) = \pi(y_i \mid x) = \exp[-(y_i - 1/d_i)^2/(2\sigma^2)]$. The measurements are mutually independed implying that the total likelihood function is a product of individual likelihoods:

$$\begin{aligned}\pi(y_1, y_2, \ldots, y_n \mid x) &\propto \pi(y_1 \mid x)\pi(y_2 \mid x)\cdots\pi(y_n \mid x)\\ &\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - 1/d_i)^2\right).\end{aligned}$$

Bayes' Formula

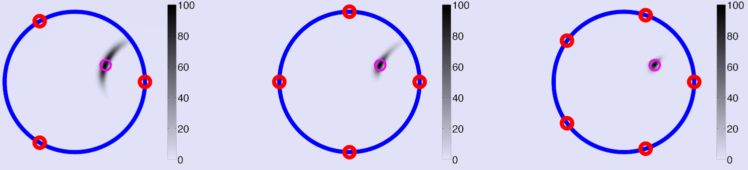
Example 1 (Matlab) continued

The prior $\pi(x)$ is, in this case, chosen to be flat (constant) in D and zero outside D. Consequently, the posterior $\pi(x \mid y)$ is given by the likelihood, i.e. $\pi(x \mid y) \propto \pi(x)\pi(y \mid x) \propto \pi(y \mid x)$. Plot the posterior with $n = 3, 4, 5$ and $\sigma = 0.1, 0.2$ using Matlab. Use percent scale with respect to the maximum value of the posterior density (MAP). Notice that the posterior is the more focused the more there are sensors and the lower is the standard deviation of the noise.

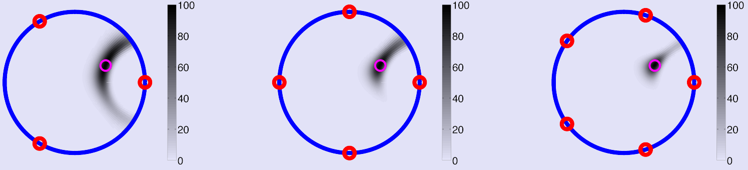
Bayes' Formula

Example 1 (Matlab) continued

Solution



3 sensors, $\sigma = 0.1$ 4 sensors, $\sigma = 0.1$ 5 sensors, $\sigma = 0.1$



3 sensors, $\sigma = 0.2$ 4 sensors, $\sigma = 0.2$ 5 sensors, $\sigma = 0.2$

Particle and sensor locations are indicated by the purple and red circles, respectively.