Probabilistic Inverse Theory

Lecture 12

W. Dębski

14.06.2023

Resolution/Error/Uncertainty analysis

A very important aspect of application of the inverse theory is a widely understood uncertainty analysis including such issue as *a posteriori* error analysis, data appraisal, sensitivity analysis, model selection, experiment planning to name a few.

Resolution/Error/Uncertainty analysis

Within probabilistic formalism all such tasks can be systematically (at least in theory) treated provided *a posteriori* pdf can be constructed and sampled. Depending on a task in hand various characteristic of the *a posteriori* pdf can be examinated.

Point-characteristics of *a posteriori* pdf

- Absolute maximum position
- Absolute maximum value
- Average (expected) value
- Covariance (resolution) mattrix
- Skewness
- Curtosis

Meta-characteristics of *a posteriori* pdf

- Evidence
- Entropy
- Relative Entropy
- Fisher Matrix

Evidence - (Z)

✦ Evidence :

$$Z[\sigma] = \int_{\mathcal{M}} \bar{\sigma}(\mathbf{m})\mathrm{d}\mathbf{m}$$

$$\sigma(\mathbf{m}) = \frac{1}{Z}\bar{\sigma}(\mathbf{m})$$

Evidence

$$\bar{\sigma}(m) = \exp\left(-\frac{m^2}{2C_p^2}\right)$$

$$Z[\sigma] = \sqrt{(2\pi C_p^2)}$$

Entropy (H)

✦ Entropy:

$$H[\sigma] = -\int_{\mathbf{M}} \ln(\sigma(\mathbf{m}))\,\sigma(\mathbf{m})\mathrm{d}\mathbf{m}$$

$$H[\sigma] = -\int_{\mathbf{M}} \ln\left(\frac{\sigma(\mathbf{m})}{\mu(\mathbf{m})}\right)\sigma(\mathbf{m})\mathrm{d}\mathbf{m}$$

Entropy cd ...

$$\sigma(m) = \frac{1}{\sqrt{(2\pi C_p^2)}} \exp\left(-\frac{m^2}{2C_p^2}\right)$$

$$\mu(m) = 1$$

$$H[\sigma] = \frac{1}{2} \ln\left(2\pi e C_p^2\right)$$

Entropy cd ...

$$\sigma(m) \;=\; \frac{1}{Z} f(\mathbf{m})\, L(\mathbf{m})$$

$$H[\sigma] \;=\; -\int_{\mathbf{M}} \ln\left(\sigma(\mathbf{m})\right)\, \sigma(\mathbf{m}) \mathrm{d}\mathbf{m}$$

$$\ln(\sigma(m)) \;=\; -\ln(Z) + \ln(f(\mathbf{m}) + \ln(L(\mathbf{m}))$$

$$H[\sigma] \;=\; H[f] + H[L] \;-\; \ln(Z)$$

Kullback-Leibler divergence measure

✦ (Relative entropy)

$$H_\psi[\sigma] = -\int_{\mathbf{M}} \ln\left(\frac{\sigma(\mathbf{m})}{\psi(\mathbf{m})}\right)\psi(\mathbf{m})\mathrm{d}\mathbf{m}$$

Relative entropy - example

$$\sigma(m) = \frac{1}{Z} f(\mathbf{m})\, L(\mathbf{m})$$

$$H_f[\sigma] = -\int_{\mathbf{M}} \ln\left(\frac{\sigma(\mathbf{m})}{f(\mathbf{m})}\right) f(\mathbf{m}) \mathrm{d}\mathbf{m}$$

$$H_\sigma[f] = -\int_{\mathbf{M}} \ln\left(\frac{f(\mathbf{m})}{\sigma(\mathbf{m})}\right) \sigma(\mathbf{m}) \mathrm{d}\mathbf{m}$$

Quantification information gain in inversion procedure

Resolution/Error/Uncertainty analysiswithin

Above discussed meta-characteristics of *a posteriori* distribution allows virtually any analysis including error (resolution, sensitivity, etc.) analysis with respect to $\mathbf{m}$ parameters, make data appraisal (extended inversion over $\mathcal{P} = \mathcal{M} \times \mathcal{D}$ space) or relation between $\mathbf{m}$ and $\mathbf{d}$ (theory selection), etc. However, in all cases a full knowledge of $\sigma(\mathbf{m})$ is required. Sometimes we cannot afford to calculate it even with very modern samplers. !!!

Resolution/Error/Uncertainty analysiswithin

We have to come back to a simpler but versatile **optimization approach** and try to incorporate to inversion an approximate and simplified error analysis.

Optimization approach - direct search for best m

Optimization approach

$$S(\mathbf{m}) = ||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m})|| + ||\mathbf{m} - \mathbf{m}^{apr}||$$

solution: search for $\mathbf{m}^{ml}$ minimizing $S(\mathbf{m})$

$$S(\mathbf{m}^{ml}) = \min$$

Inversion errors - Bayesian orientated

$$||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{est})||_D + ||\mathbf{m}^{est} - \mathbf{m}^{apr}||_M = \min$$

$$\begin{aligned} ||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{est}) + \epsilon^{obs} + \epsilon^{th}(\mathbf{m})||_D &+ \\ ||\mathbf{m}^{est} - \mathbf{m}^{apr} + \epsilon^{apr}||_M &= \min \end{aligned}$$

$$\Downarrow$$

$$\mathbf{m}^{est}_{ijk} = \mathbf{m}^{est}\left(\epsilon_i^{obs}, \epsilon_j^{th}(\mathbf{m}^{est}), \epsilon_k^{apr}\right)$$

$$\Downarrow$$

$$\sigma(\mathbf{m}_{ijk}) = \frac{N_{ijk}}{N_{all}}$$

Maksimum Likelihood Method - frequentist orientated

In case when only observational errors are considered (e.g. in case when data comes from repeated measurements of the same physical parmeter and are treated as a stochastic ($\mathbf{d}^i$) parameter

$$||\mathbf{d}^{true} - \mathbf{d}^{th}(\mathbf{m}^{est}) + \epsilon^{obs}||_D + ||\mathbf{m}^{est} - \mathbf{m}^{apr}||_M = \min$$

$$||\mathbf{d}^{i} - \mathbf{d}^{th}(\mathbf{m}^{est})||_D + ||\mathbf{m}^{est} - \mathbf{m}^{apr}||_M = \min$$

One can construct

$$P(\mathbf{d}^i|\mathbf{m}) = \exp\left(-||\mathbf{d}^i - \mathbf{d}^{th}(\mathbf{m})||\right)$$

Maksimum Likelihood Method

Double interpretation of $P(\mathbf{d}^i|\mathbf{m})$

- Data errors (noise) function (conditioned on $\mathbf{m}$

$$N(\mathbf{d}^i) = P(\mathbf{d}^i|\mathbf{m})$$

- Likelihood function (conditional probability)

$$\Psi(\mathbf{m}) \;=\; P(\mathbf{d}^i|\mathbf{m})$$

Solution

$$\Psi(\mathbf{m}^{est}) \;=\; \max$$

is a model which most likelihood reproduce all $\mathbf{d}^i$

How to estimate errors for MLL approach ?

Let us consider $N(X|\mathbf{m})$ distribution where X represents stochastic variable and $\mathbf{m}$ is a set of "fixed" parameters.

Let us define score:

$$S(X) \ = \ \frac{\partial}{\partial \mathbf{m}} \log N(X|\mathbf{m})$$

$$< S >_X = \int\limits_X S(X) N(X) dX = \frac{\partial}{\partial \mathbf{m}} \int\limits_X N(X) dX = 0$$

Fisher information

$$\mathcal{I}(\mathbf{m}) = \left\langle \left(\frac{\partial}{\partial \mathbf{m}} \log N(X|\mathbf{m}) \right)^2 \right\rangle$$

$$\mathcal{I}(\mathbf{m}) = \int_X \left(\frac{\partial}{\partial \mathbf{m}} \log N(X|\mathbf{m}) \right)^2 N(X) dX$$

$\mathcal{I}(\mathbf{m})$ express amount of information about unknown $\mathbf{m}$ carried out by random variable X

Cramer-Rao bound

Inverse of Fisher information is a lower bound on the variance of unbiased estimator of $\mathbf{m}$

$$\langle \bar{\mathbf{m}}(X) - \mathbf{m} \rangle = \int (\bar{\mathbf{m}}(X) - \mathbf{m}) N(X) = 0$$

$$\frac{\partial}{\partial \mathbf{m}} \int (\bar{\mathbf{m}}(X) - \mathbf{m}) N(X) = 0$$

$$\int (\bar{\mathbf{m}} - \mathbf{m})^2 N(X) dX \cdot \int \left(\frac{\partial log N}{\partial \mathbf{m}} \right)^2 N(X) dX \ \geq 1$$

End of Lecture

Continuation:

Hands on Advanced Inverse Issues