Theoretical background SFIT4

Mathias Palm

Institute of Environmental Physics
Universität Bremen

Tsukuba, June 2013

Statement of the problem

The forward model

- The measured quantity $\vec{y}$ is a result of a physical process.
- Creation of the measured quantity can usually be modeled (the Forward Model F):

$$\vec{y} = F(\vec{x}, \vec{b}_0, \vec{b}_1, \ldots) + \vec{\epsilon}$$

- **Wanted:** The inverse F^{-1}.
- But: F^{-1} may not exist, may not be unique or difficult to evaluate
- Work around: Solve for

$$\min_{x, b_0, b_1, \ldots} \left(\hat{\vec{y}} - F(\vec{x}, \vec{b}_0, \vec{b}_1, \ldots) \right) \tag{1}$$

- **Another problem** no or many solutions may exist

Statement of the problem

Possible solutions

Thikonov-Phillips-regularization Solution is x which minimises the functional

$$\hat{x} = \min_x || F(x) - y || + \lambda || L(x) ||$$

$|| \cdot ||$ can be any norm

Optimal esimation solution is conditional probability distribution $p(x|y)$ given by (Bayes theorem)

$$p(y) \quad \underbrace{p(x|y)}_{\text{a posteriory}} \quad = \quad \underbrace{p(y|x)}_{\text{likelihood}} \; \underbrace{p(x)}_{\text{apriori}}$$

Formulation of the forward model as probability density

The forward model is linear:

$$F(\vec{x}) = K\vec{x} + \vec{\epsilon} \tag{2}$$

The noise $\vec{\epsilon}$ is Gaussian with mean zero and white. Hence:

$$p(y|x) = C_1 \exp\left(-(K(x) - y)^T S_\epsilon^{-1} (K(x) - y)\right) \tag{3}$$

Notation: $p(x|y)$ is probability of x given y, i.e. conditional probability

Formulation of the a priory probability density

Formulation

$$p(x) = C_2 \exp\left(-(x - x_A)^T S_A^{-1} (x - x_A)\right) \tag{4}$$

- The a priory contains knowledge of the state being measured. It is assumed that the state is distributed like a Gaussian.
- This is problematic in real world applications.
- Chosen for mathematical tractability.

A posterior distribution

- Gaussian shape (because a priori and likelihood are Gaussian)

$$
\begin{aligned}
p(x|y) &= C_3 \exp\left(-(x-\hat{x})^T \hat{S}^{-1}(x-\hat{x})\right) \\
\hat{x} &= (S_A^{-1} + K^T S_\epsilon^{-1} K)^{-1} \\
& \quad K^T S_\epsilon^{-1}[y - F(x_i)] - S_A^{-1}(x_i - x_A) \\
&= (S_A^{-1} + K^T S_\epsilon^{-1} K)^{-1} K^T S_\epsilon^{-1}(y - K x_A) \\
\hat{S} &= (S_A^{-1} + K^T S_\epsilon^{-1} K)
\end{aligned}
$$

- The a posterior has one **MODE.**
- mode = median = expected value

Calculation is easy and not time consuming

Calculating the mode of the a posterior distribution

The forward model is solved if it is

linear: calculation in one step

weakly non-linear: iterative using the Gauss - Newton - Algorithm
-> approximation of the newton algorithm

strongly non-linear: iterative using the Levenberg Marquardt algorithm
-> mixture of steepest descend and Gauss-Newton