DRACULA-ADVANCED RETRIEVAL TOOL FOR ATMOSPHERIC REMOTE SENSING

A. Doicu, F. Schreier, S. Hilgers, A. von Bargen, and B. Aberle

DLR — German Aerospace Center, Remote Sensing Technology Institute,
Oberpfaffenhofen, 82234 Weßling, Germany

ABSTRACT

In this paper we present a regularization tool for solving inverse problems arising in atmospheric remote sensing. The tool incorporates several direct and iterative regularization methods based on different basic principles. To deal with model parameters that introduce uncertainties which cannot be ignored we use the marginalizing technique. The tool also includes a postprocessing part for characterizing the solution accuracy.

Key words: retrieval, regularization.

1. INTRODUCTION

The main reason for the design of a regularization tool including several inversion methods (based on different basic principles) is that for a specific application, we may select the optimal approach from the point of view of accuracy and efficiency. On the other hand, a regularization method is characterized by some parameters which control the inversion process. These control parameters include: the penalty term (a priori state vector, type of regularization matrix, strength of regularization) and the termination criterion. For the operational usage of an on-line processor, the control parameters must be determined in advanced, and this process can be a result of the intercomparison of several regularization methods.

The aDvanced Retrieval of the Atmosphere with Constrained and Unconstrained Least Squares Algorithms (DRACULA) tool incorporates several regularization methods, and communicates with the forward model via a subroutine which computes the forward model and the Jacobian at the current iterate. The problem under consideration tool for solving the nonlinear problem (1) includes a preprocessing, a processing and a postprocessing part.

2. PREPROCESSING

The preprocessing part of the regularization tool consists of two steps: marginalizing and prewhitening. In the marginalizing step, the noise covariance matrix is recalculation accounting of model uncertainties, that is,

$$\hat{\Sigma}_\delta = \Sigma_\delta + K_b C_b K_b^T,$$

where

$$K_b = \frac{\partial F}{\partial b} (x_\delta, b_b)$$

is the Jacobian matrix corresponding to the vector of auxiliary parameters b and C_b is the covariance matrix of model uncertainties.

In the prewhitening step, the data model with covariance matrix C_δ is transformed into a data model with white noise covariance matrix. For this purpose, we consider the singular value decomposition C_δ = U_δ Σ_δ U_δ^T, where Σ_δ = diag(η_δ), and define the ‘equivalent’ white noise variance σ_δ^2 = (1/m) Σ_k=1^m η_δ^2 and the normalized noise covariance matrix C_δ = (1/σ_δ^2) ̂C_δ. Multiplying equation (1) by ̂C_δ^{-1/2} , where ̂C_δ^{-1/2} = U_δ ̂Σ_δ^{-1/2} U_δ^T and Σ_δ = (1/σ_δ^2) ̂Σ_δ, and considering the change of variables ̂C_δ^{-1/2} y^d → y^d, ̂C_δ^{-1/2} F → F, and ̂C_δ^{-1/2} → δ, we obtain the data model (1) with E{δ} = 0 and C_δ = E{δδ^T} = σ_δ^2 I_m.

3. PROCESSING

The processing part consists of several regularization method for solving the nonlinear ill-posed problem. The following inversion methods are included in the regularization tool:

1. Tikhonov regularization computes the regularized solution x_δ^E by minimizing the objective function

$$F(x) = \frac{1}{2} \left( ||F(x) - y^d||^2 + \alpha ||L(x - x_a)||^2 \right).$$
When the Gauss–Newton method is used as optimization tool, the method is equivalent to the solution of a sequence of linear problems at each Newton step: Assuming a linearization of the forward model $F$ about the current iterate $x_{k\alpha}$, the linear problem

$$K_{k\alpha} (x - x_{k}) = y_{k}^{\delta},$$

with

$$y_{k}^{\delta} = y^{\delta} - F (x_{k\alpha}^{\delta}) + K_{k\alpha} (x_{k\alpha}^{\delta} - x_{k}),$$

is solved by means of Tikhonov regularization with a penalty term of the form

$$\Omega (x) = \|L (x - x_{k})\|.$$ 

The regularization parameter can be chosen by using a priori selection criteria (expected error estimation method, statistical information) or error–free parameter selection criteria (generalized cross validation, L–curve method, maximum likelihood estimation, quasi–optimality criterion).

2. Iterated Tikhonov regularization solves the linearized problem (2) by means of an iterative scheme. The first iteration step is the step of the ordinary method, while at the iteration step $j$, we evaluate the discrepancy of the linear equation and formulate a new linear equation in terms of the improved solution step and the discrepancy data vector (as the right-hand side), i.e.,

$$K_{k\alpha} (x - x_{k\alpha j}) = y_{k}^{\delta} - K_{k\alpha} (x_{k\alpha j}^{\delta} - x_{k}).$$

This equation is again solved by means of Tikhonov regularization and it is apparent that this method consists of an outer loop - the Newton iteration for the nonlinear problem - and an inner loop - the iterated Tikhonov regularization for the linearized problem. The inner loop involves $p$ steps, where $p$ is order of the iterated Tikhonov regularization. The order of the iterated Tikhonov regularization is a control parameter of the algorithm and must be determined in advance.

3. Newton-type methods with a priori information solves the linearized problem

$$K_{k} (x - x_{k}) = y_{k}^{\delta},$$

by using the iteration

$$x_{k+1}^{\delta} = x_{k} + g_{\alpha} \left[ (L^{T} L)^{-1} K_{k}^{T} K_{k} \right] (L^{T} L)^{-1} K_{k}^{T} y_{k}^{\delta}.$$ 

The parameter dependent family of real valued functions $g_{\alpha}$ characterizes the linear regularization method and may corresponds to

1. Tikhonov regularization,
2. Landweber iteration,
3. iterated Tikhonov regularization of fixed order,
4. $\nu$-methods or
5. conjugate gradient.

$\{\alpha_{k}\}$ is an a priori (independently of the noise level) fixed sequence of parameters, satisfying the relations

$$\alpha_{k} > 0, \quad 1 \leq \frac{\alpha_{k}}{\alpha_{k+1}} \leq c, \quad \lim_{k \to \infty} \alpha_{k} = 0,$$

for some $c > 1$. The iterative process is stopped accordingly to the discrepancy principle, instead of requiring the convergence of iterates. The widely used Newton-type method with a priori information is the iteratively regularized Gauss-Newton method. This approach is characterized by the choice $g_{\alpha} (\lambda) = 1/(\lambda + \alpha)$ and the iteration is identical to that of the Tikhonov regularization with a variable regularization parameter.

4. Newton-type methods without a priori information use filter functions for regularizing the linearized equations

$$K_{k} (x - x_{k}^{\delta}) = r_{k}^{\delta},$$

where

$$r_{k}^{\delta} = y^{\delta} - F (x_{k}^{\delta}).$$

The resulting iteration is

$$x_{k+1}^{\delta} = x_{k}^{\delta} + g_{\alpha} \left[ (L^{T} L)^{-1} K_{k}^{T} K_{k} \right] (L^{T} L)^{-1} K_{k}^{T} r_{k}^{\delta},$$

and as for Newton-type methods with a priori information, the parameters $\alpha_{k}$ are a priori chosen and satisfy (3). The representative approach which belongs to this class of regularization techniques is the Levenberg-Marquardt method, characterized by the choice $g_{\alpha} (\lambda) = 1/(\lambda + \alpha)$.

The regularization matrix $L$ is used as penalty term for Tikhonov type methods and as preconditioner for iterative regularization methods of Newton type. $L$ can be chosen as discrete approximations to derivative operators, as Sobolev norm including several derivative approximations, or by using statistical information.

4. POSTPROCESSING

The postprocessing part consists of an error analysis, which is performed for Tikhonov type methods. Essentially, we compute the mean square error matrix

$$S_{\alpha} = (I_{n} - A_{\alpha}) (x^{\delta} - x_{\alpha}) (x^{\delta} - x_{\alpha})^{T} (I_{n} - A_{\alpha})^{T} + \sigma^{2} K_{\alpha}^{T} K_{\alpha}^{T}$$

to quantify the dispersion of the regularized solution $x_{\alpha}^{\delta}$ about the true solution $x^{\delta}$. The one-rank matrix $(x^{\delta} - x_{\alpha}) (x^{\delta} - x_{\alpha})^{T}$ can be approximated by $(x_{\alpha}^{\delta} - x_{\alpha}) (x_{\alpha}^{\delta} - x_{\alpha})^{T}$, or by $(\sigma^{2}/\alpha) (L^{T} L)^{-1}$, in which case

$$S_{\alpha} \approx \sigma^{2} (K_{\alpha}^{T} K_{\alpha} + \alpha L^{T} L)^{-1}.$$
A linearized error analysis can be performed when the sequence of iterates \( \{ x^\delta_{k\alpha} \} \) converges, the forward model can be linearized about \( x^\delta_{\alpha} \), and the data error model is correct.

The linearity assumption is verified at the boundary of a confidence region for the solution. For this purpose we consider a singular value decomposition of the means square error matrix

\[
S_{\alpha} = V \Sigma V^T
\]

and define the normalized error patterns \( s_k \) for \( S_{\alpha} \) from the decomposition

\[
V \Sigma^{1/2} = [s_1, ..., s_n].
\]

We then compute the values of the objective function \( F \) at selected points on the boundary of this region, e.g., \( F (x^\delta_{\alpha} + s_k) \) for \( k = 1, ..., n \), and if our linearity assumption is valid, these values should differ only slightly from the approximate values

\[
M (s_k) = F (x^\delta_{\alpha}) + \frac{1}{2} s_k^T (K^T_{\alpha} K_{\alpha} + \alpha L^T L) s_k.
\]

Alternatively, we may consider the linearization error

\[
R (x) = F (x) - F (x^\delta_{\alpha}) + K_{\alpha} (x - x^\delta_{\alpha}),
\]

and check if the condition

\[
\| R (x^\delta_{\alpha} \pm s_k) \|^2 \leq \sigma^2,
\]

holds true for all \( k = 1, ..., n \).

The validity of the data error model is perhaps the most important problem of an error analysis. The presence of the systematic errors introduce an additional bias in the solution, and to account of this type of errors, we replace the data error \( \delta_y \) by an equivalent white noise \( \delta_e \) such that

\[
E \left\{ \| \delta_e \|^2 \right\} = E \left\{ \| \delta_y \|^2 \right\}.
\]

The variance of the white noise \( \delta_e \) is estimated by

\[
\sigma_e^2 \approx \frac{1}{m - n} \| F (x^\delta_{\alpha}) - y^\delta \|^2,
\]

in the limit \( \alpha \to 0 \). This equivalence enables an orientative error analysis with the white noise covariance matrix \( C_{\delta_e} = \sigma_e^2 I_m \).

5. NUMERICAL SIMULATIONS

In Figures 1 and 2 we plot the retrieval results computed by Tikhonov regularization (TR), iteratively regularized Gauss–Newton method (IRGNM), Newton–Landweber iteration (NLI) and Levenberg–Marquardt method (LMM). The results demonstrate that in the case of Tikhonov regularization, the selection of the a priori state vector, of the regularization matrix and of the strength of regularization is optimal.
6. CONCLUSIONS

A regularization tool for atmospheric remote sensing has been presented. The tool is devoted to the solution of multi-parameter and bound-constrained inversion problems and incorporates direct and iterative regularization methods (Tikhonov regularization, iteratively regularized Gauss–Newton method, Levenberg–Marquardt method, etc.). Several parameter-choice methods including the L-curve method, the generalized cross-validation approach, the unbiased predictive risk estimator method, the minimum bound method and the noise error criterium are implemented in the code. These methods guarantee an optimal choice of the regularization parameter. Diagnostic tools (the discrete Picard condition and the histories of iterates and residuals) offer additional information about the iterative process. The error analysis is performed in a semi-stochastic setting and consists in the computation of the mean square error matrix and the estimation of the smoothing and noise errors. All these features lead to an efficient and robust inversion algorithm for atmospheric remote sensing.