ATMOS 6500, Fall 2019
Numerical Weather Prediction
Dr. Zhaoxia Pu

Lab Assignment V: Four-dimensional variational data assimilation (4DVAR)

Assigned Wednesday, October 30, 2019
Due Monday, November 4, 2019

Objectives:

1) Get familiar with adjoint and tangent linear models; 2) Understand the components to build up a 4DVAR system; 3) Learn 4DVAR data assimilation procedures; 4) Practice data assimilation with 4DVAR

Instructions:

1. The Lorenz equations

We consider a four-dimensional variational data assimilation scheme applied to the Lorenz equations, a simple dynamical model with chaotic behavior. The Lorenz equations are given by a nonlinear system:

$$\frac{dx}{dt} = -\sigma(x - y)$$
$$\frac{dy}{dt} = \rho x - y - xz$$
$$\frac{dz}{dx} = xy - \beta z$$

where *x=x(t), y=y(t), z=z(t)* and sigma, rho, beta are parameters, which in these experiments are chosen to have the values 10, 28 and 8/3 respectively. The system is discretized using a second order Runge-Kutta method.

2. 4DVAR

You will practice the 4DVAR method using the sample Matlab program I provided to you. The true state and the background state are defined in the program. You can compare the results from different experiments with different observation frequencies and accuracies.

2.1 Introduction

The 4D-Var schemes in these programs have only an observation term, so minimize a function of the form

$$J = \frac{1}{2}\sum_{i=0}^{n}(y_i - H_i(x_i))^T R^{-1}(y_i - H_i(x_i))$$

A 4D-Var scheme minimizes this cost function with the use of the nonlinear model and the corresponding adjoint. I provided you the sample programs for 4DVAR. The components of the programs are listed as follows:

lorenz4d.m	----- Top level routine for full 4D-Var
modeuler.m	----- Nonlinear model for Lorenz system
modeuler_tl.m	----- Tangent linear model
modeuler_adj.m	----- Adjoint model
test_tl.m	----- Test tangent linear model
test_adj.m	-----Test adjoint model

2.2 Test routines - Building a 4D-Var system

When building a 4D-Var system, there are standard ways of testing the various components before they can be used for assimilation.

2.2.1 Test a tangent linear model

Suppose that M is a nonlinear model and L is a tangent linear model. Then, for a small perturbation $\gamma\delta x$, we have

$$M(x+\gamma\delta x) - M(x) \approx L(x)\gamma\delta x$$

Hence if we plot the relative error

$$E_R = \frac{M(x+\gamma\delta x) - M(x)}{L(x)\gamma\delta x} - 1$$

As $\gamma \to 0$, we should find that $E_R \to 0$.

Exercise: Use the routine *test_tl* to plot the relative error. Introduce an error into the tangent linear code *modeuler_tl* and see what the impact is.

2.2.2 Test an adjoint model

For a linear model L and its adjoint L^* we have the identity

$$< L\delta x, L\delta x > = < \delta x, L^*\delta x >$$

for any inner product <,> and perturbation δx. This can be applied to test whether the adjoint is coded correctly.

Exercise: Apply the routine *test_adj* to test the adjoint code. Introduce an error into the adjoint code *modeuler_adj* and see what the impact is.

2.3 Assimilation program

The routines used to run assimilation experiments is *lorenz4d* for the 4D-Var. The menu options you must specify, with some suggested values, are

Initial values of x,y,z:	0.0--5.0
Assimilation period (in seconds):	0--10
Forecast period (in seconds):	Any
Time step (in seconds):	0.0--0.05
Frequency of observations (in time steps):	Any
Noise on observations:	Variance = 0 -- 2
Convergence criteria:	Default values are given

Note that the time step must be a divisor of your total time, so values such as 0.02, 0.025, 0.05 work well. The output of the program is the fields of x and z, the errors in x and z and the convergence of the cost function and its gradient. The final norm of the gradient is also output in the Matlab command window.

2.4 Suggested exercises

Start with the conditions

Truth=(1.0,1.0,1.0)
First guess=(1.2,1.2,1.2)
Assimilation period = 2
Forecast period = 3
Time step = 0.05
Frequency of observations = 2

Play with the code, change the parameters above and try to answer the following questions:

- Compared with the case when we only have very few accurate observations, would it be better (or worse) to have more observations which are less accurate? Address this problem by comparing both accuracy of the analysis and rate of the convergence.

- Would it be better to have a long assimilation window with few observations or a short assimilation window with more observations? Does this result depend on the error variances the observations?

- How does the rate of convergence change if the first guess is moved closer to or away from the truth?

- The convergence depends on the iteration number and given tolerance. Perform experiments to investigate how the accuracy of analysis changes with iteration numbers and given tolerance.

- Compared different assimilation period for a given total assimilation and forecast time, how does the convergence vary?

Lab Report:

Based on your experiments, write a report to summarize your results (2 page minimum, 12pt Times New Roman; doubled space). You can attach figures to support your conclusion.