

数值模式初值的敏感性程度对四维同化的影响 ——基于 Lorenz 系统的研究*

郜吉东^① 丑纪范

(兰州大学大气科学系, 兰州, 730000)

摘 要

用著名的 Lorenz 系统作了共轭变分同化的数值试验。发现随着模式对初值敏感性程度的增加,用这种方法得到和模式相协调的初始场愈来愈困难,直到某些情况下的完全失败。这表明四维同化和可预报期限是联系在一起的。另一方面,随着方程不精确程度的增加,变分同化的效果愈来愈差,直到所做的预报无任何意义可言。如果在做变分同化的同时对模式参数也进行反演,就可使得基于 Lorenz 系统所做的预报效果大大提高。

关键词: 变分同化, Lorenz 系统, 共轭方程。

1 引 言

共轭变分同化的思想是由 Lewis 和 Derber (1985)^[1], Talagrand 和 Courtier^[2]等的开创性工作开始实现的。实质上它是最优化方法在大气或海洋方面的具体应用。因其理论上的完善性,能非常自然地把观测资料和数值模式巧妙地结合在一起,并且易适用于对非定时资料的同化,而日益引起人们的重视。国际上的几个气象中心,如欧洲的 ECMWF, 美国的 NCAR 等都在着手把这种方法业务化,并且取得了一些重要进展。如同其它同化方法,这种方法也有其局限性,例如,用此种方法得到的极小值不一定是全局极小点^[3,4],又如通过仅仅调整初值求得最佳的模式解,假定数值模式是准确无误差的等。

另一个必须注意的问题是,当数值模式是高度非线性的。如存在混沌解时的情形,这时有共轭变分同化方法就存在着比较大的困难,因模式对初值高度敏感,共轭方程在资料同化中的作用变得微不足道,资料同化的意义已不明显,甚至没有必要。对于 Lorenz 系统,很多人已做过深入的研究^[4],在某种程度上能够代表实际大气的一些特征,通过取不同的参数,可以呈现出不同的对流态,并且其共轭系统非常容易导出。所以本文先用 Lorenz 系统作了这方面的数值试验。

2 变分同化的基本原理

变分同化的基本思想是定义一个目标泛函 J 来度量同化时段 (t_0, t_N) 上模式解与观测

* 1993 年 7 月 31 日收到原稿, 1994 年 7 月 12 日收到修改稿。受国家青年自然科学基金及教委高等学校博士点专项科研基金资助。

^① 现在北京气象学院大气科学系工作。

之间的某种“距离”，然后以同化模式的控制方程组为约束，用最优化方法反演模式的初值 u_0 ，使 J 达到极小点。这种方法的关键是要求得目标泛函关于控制变量 u_0 的梯度 $\partial J / \partial u_0$ ，它将由同化模式的共轭方程向后积分（从 $t_N \rightarrow t_0$ ）而得到。整个问题可用如下数学模型来描述^[5]：

$$\begin{cases} \frac{du}{dt} = H(u) & (1) \\ J(u) = \sum_{j=1}^N \langle u(t_j) - \hat{u}(t_j), u(t_j) - \hat{u}(t_j) \rangle & (2) \end{cases}$$

所谓变分同化指的是在约束(1)下，式(2)的极小值。这里式(1)代表数值模式， H 为由 Hilbert 空间 $\Omega \rightarrow \Omega$ 的非线性算子。式(2)是所定义的目标泛函， \langle, \rangle 代表内积（定义在 Ω 上）。式(1)的一切时刻的值由初值 $u_0 = u(t_0)$ 唯一决定。因而一旦 $u(t_0)$ 求出，由约束(1)自然求出 $u(t_j)$ ， $j = 1, 2, \dots, N$ 。此时可把式(2)变为无约束求极小。并且可以计算出对 $u(t_0)$ 的梯度。

现在推导 $J(u)$ 关于 $u(t_0)$ 的梯度，对式(2)求变分：

$$\delta J = 2 \sum_{j=0}^N \langle u(t_j) - \hat{u}(t_j), \delta u(t_j) \rangle \quad (3)$$

由式(1)得到：

$$\frac{d\delta u}{dt} = A(t) \cdot \delta u \quad (4)$$

其中， $A(t)$ 为 $H(u)$ 的雅可比算子（导算子），它是一个线性算子。

由式(4)，类似一般的常微分方程，对于一切 $t, t', \delta u(t), \delta u'(t) \in \Omega$ ，存在线性算子 $\mathcal{L}(t, t')$ ，使

$$\delta u(t) = \mathcal{L}(t, t') \cdot \delta u(t') \quad (5)$$

成立，因而由式(3)，得：

$$\delta J = 2 \cdot \sum_{j=1}^N \langle u(t_j) - \hat{u}(t_j), \mathcal{L}(t_j, t_0) \delta u(t_0) \rangle \quad (6)$$

设 $\mathcal{L}^*(t, t')$ 是 $\mathcal{L}(t, t')$ 的伴随算子，则由 Hilbert 算子理论^[6]，

$$\delta J = 2 \sum_{j=0}^N \langle \mathcal{L}^*(t_j, t_0)(u(t_j) - \hat{u}(t_j)), \delta u(t_0) \rangle \quad (7)$$

由此得到 J 关于 $u(t_0)$ 的梯度为：

$$\nabla J = 2 \sum_{j=0}^N \mathcal{L}^*(t_j, t_0)(u(t_j) - \hat{u}(t_j)) \quad (8)$$

可以证明 $\mathcal{L}^*(t_j, t_0)$ 是伴随方程

$$\frac{d\delta u}{dt} = -A^*(t) \cdot \delta u^* \quad (9)$$

的预解式。

有了梯度(8)的计算表达式之后，用某类最优控制算法，如共轭梯度法、最速下降法等，可以计算式(1)、(2)的条件极小。整个算法也可称之为对初值的反演。

3 Lorenz 模型及其共轭模式

Lorenz 模型是 Lorenz 和 Saltzman 1963 年在研究流体有限振幅对流时提出的非线性谱模式,其形式为:

$$\begin{cases} \frac{dx}{dt} = -\sigma X + \sigma Y \\ \frac{dY}{dt} = -XZ + rX - Y \\ \frac{dZ}{dt} = XY - bZ \end{cases} \quad (10)$$

这里 σ, r 和 b 分别是 Prantl 数, Rayleigh 数及与对流尺度相联系的参数。本文中根据 Lorenz 1963 年的研究取为 $\sigma = 10, b = 8/3, r$ 根据不同的试验给定。

按照上节,易导出式(10)的切线性方程的共轭方程为:

$$\begin{aligned} -\frac{d\delta X^*}{dt} &= -\sigma\delta X^* + (-Z + r)\delta Y^* + Y\delta Z^* - \frac{\partial G}{\partial X} \\ -\frac{d\delta Y^*}{dt} &= \sigma\delta X^* - \delta Y^* + X\delta Z^* - \frac{\partial G}{\partial Y} \\ -\frac{d\delta Z^*}{dt} &= -X\delta Y^* - b\delta Z^* - \frac{\partial G}{\partial Z} \end{aligned} \quad (11)$$

其中 $(\delta X^*, \delta Y^*, \delta Z^*)$ 是 $(\delta X, \delta Y, \delta Z)^T$ 的共轭矢量。

$$G(X, Y, Z, t) = (X - X_{obs})^2 + (Y - Y_{obs})^2 + (Z - Z_{obs})^2 \quad (12)$$

是目标泛函中正则函数。目标泛函 J 可表示为:

$$J = \sum_{t=0}^N [(X - X_{obs})^2 + (Y - Y_{obs})^2 + (Z - Z_{obs})^2] \quad (13)$$

这里 $X_{obs}, Y_{obs}, Z_{obs}$ 是和模式变量相应的“观测资料”。

式(10), (11)离散化后,时间步长 $\Delta t = 0.01$, 控制试验(Control run)积分 1600 步,前 800 步每隔 20 步存放一次积分值作为变分同化(即反演初值)所用的“观测资料”,后 800 步的积分结果作为试验预报好坏的评估。选用 Beale-Powell 共轭梯度算法来极小化目标泛函,它是一种效率较高的共轭梯度算法^[7]。整个计算过程是先给出 X_0, Y_0, Z_0 的第一猜值,通过向前积分基本模式(10)和向后积分共轭型(11)求得 $(\frac{\partial J}{\partial X_0}, \frac{\partial J}{\partial Y_0}, \frac{\partial J}{\partial Z_0})$, 看这些计算值的范数是否小到满足一定的收敛条件。若不满足,用共轭梯度算法寻求新的 X_0, Y_0, Z_0 , 再重复上述步骤;若是已满足了收敛条件,则认为得到了最优的初值 X_0, Y_0, Z_0 , 在 (t_0, t_N) 上积分,积过 t_N 时刻则作出了未来的预报。若要同时反演模式参数,则把模式参数也作为控制变量,详见文献[8]。

4 数值试验

4.1 对于 Lorenz 系统,不同类型的对流态反演难易的差异

无论是用最速下降法还是用共轭梯度法求解,都希望目标泛函的形式是二次型的,并且是凸的,这样整个下降算法才能够收敛。这两个性质保持得越好,则收敛速度越快。

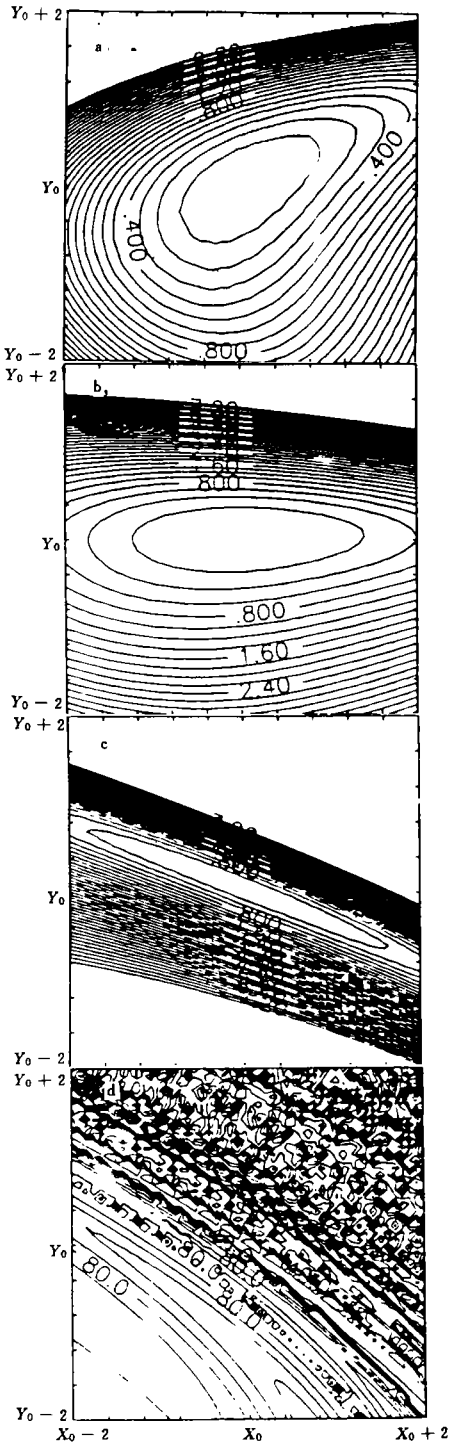


图1 r 取不同值时目标泛函等值线形状图
(a. $r = 2$; b. $r = 23$; c. $r = 26$; d. $r = 28$)

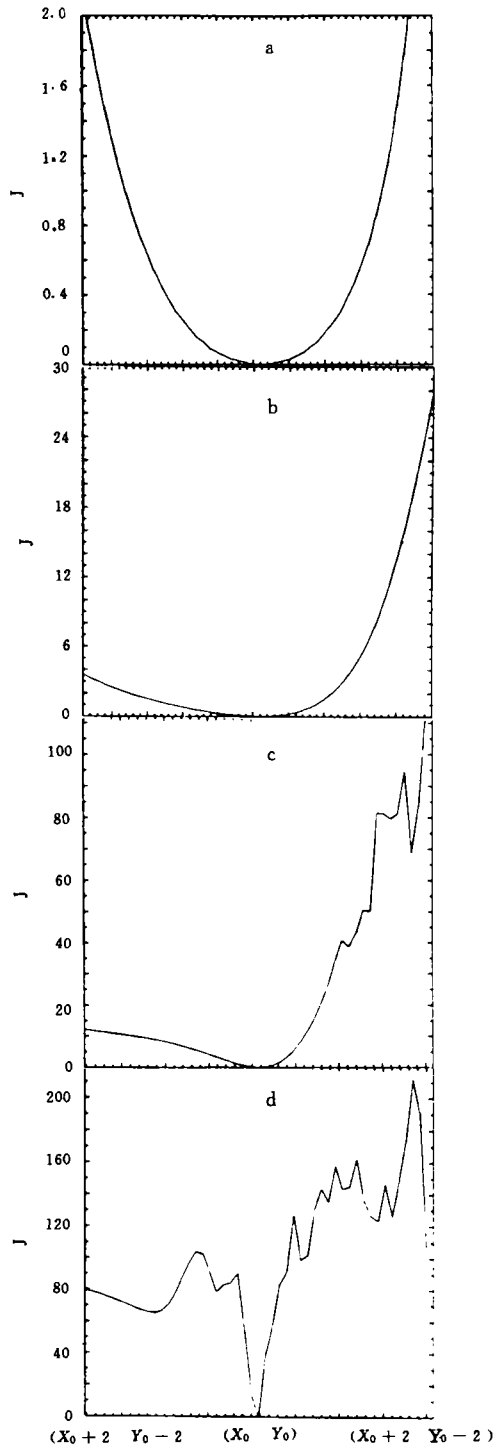


图2 和图1相对应的经过两点 $(X_0 - 2, Y_0 - 2, Z_0)$ 和 $(X_0 + 2, Y_0 + 2, Z_0)$ 的目标泛函值变化曲线

图 1 给出了 r 取四种不同的参数的情况下,对相空间中某一固定点 $(X_0, Y_0, Z_0) = (-5.92, -5.90, 24.0)$,当第一猜测值范围在 $X_0 - 2 < X'_0 < X_0 + 2, Y_0 - 2 < Y'_0 < Y_0 + 2, Z'_0 = Z_0$ 时目标泛函 J 的等值线形状。从图 1a—图 1d, 参数 r 逐渐变大,对流渐变为混沌的状态,目标泛函等值线的形状由原来的长短轴较为接近的椭圆,变为在 X 方向长轴很长,扁平的椭圆。到 $r = 28$ 时,目标泛函的等值线变得非常复杂。为说明问题方便,图 2a—2d 给出了和图 1a—1d 对应的经过 $(X_0 - 2, Y_0 - 2, Z_0)$ 到 $(X_0 + 2, Y_0 + 2, Z_0)$ 两点

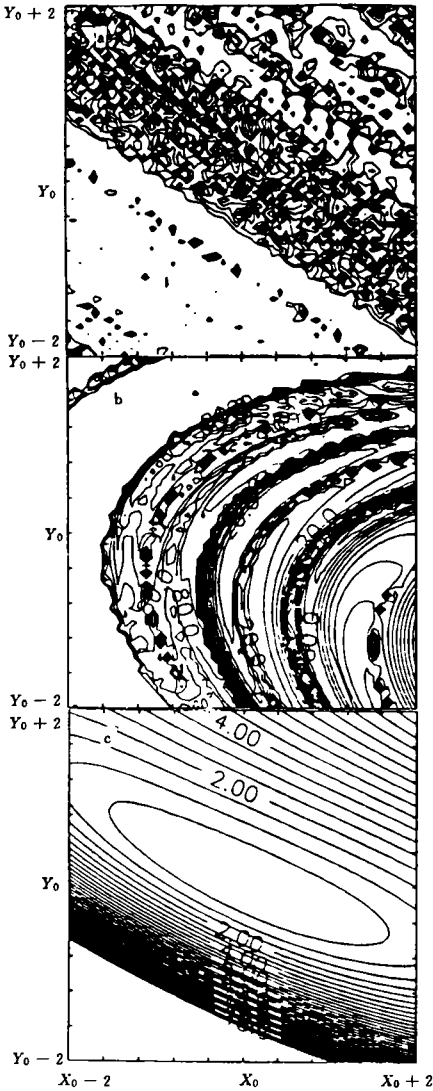


图 3 相空间中不同位置目标泛函等值线形状
 (a, $(-0.17, 0.84, 19.95)$ b, $(6.57, 8.89, 22.31)$;
 $(-9.42, -9.34, 28.31)$)

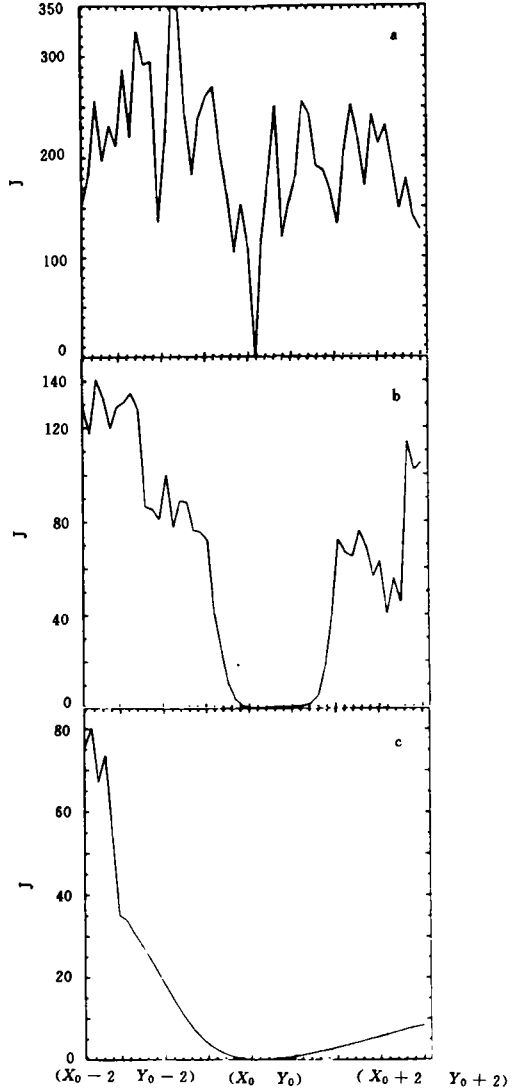


图 4 和图 3 相对应的经过两点 $(X_0 - 2, Y_0 - 2, Z_0)$ 和 $(X_0 + 2, Y_0 + 2, Z_0)$ 的目标泛函值变化曲线

连线之间的目标泛函曲线变化图。图 2a 和图 2b 等值线比较光滑。表明从其上任一点开始迭代的点,都将下降到全局极小点。尽管由图 2a 到图 2b 下降速度变慢,图 2c—d 曲线出现很多折点,开始出现局部极小点,而且收敛速度变得更慢,这样如果第一猜值取得不合适,很可能收敛到局部极小点。对于图 2d,无论用任何效率高的最优控制技术,恐怕都难达到全局极小点。

对于不同的参数 r , Lorenz 系统分别对应着不同的对流态。随着 r 的增大,流型由对流态变到混沌态,非线性的作用愈来愈强,模式对初值也愈来愈敏感,初值稍有不同,解的轨迹完全两样。因而要反演得到使模式解和观测相适应的初值,求解问题的病态程度愈来愈严重,困难愈来愈大。

下面的数值试验表明,也可把目标泛函等值线形状作为衡量模式对初值敏感程度的一个标准。和图 1a—d 相应的数值试验表明,目标泛函等值线椭圆的长短轴长度愈接近,模式对初值愈不敏感,初值愈容易反演;椭圆愈扁平,模式对初值愈敏感,初值反演的难度愈大;当目标泛函等值线形状毫无规则时,则模式对初值高度敏感,反演已无任何意义。

4.2 在混沌态($r=28$)时相空间中不同位置可反演程度的差异

在参数 $r=28$ 时,图 1d 表明,点 $(X_0, Y_0, Z_0) = (-5.92, -5.90, 24.0)$ 附近,目标泛函等值线形状非常复杂。但相空间中其它的点,却不一定有这么复杂,可再取另外三点 $(-9.42, -9.34, 28.31)$, $(-0.17, 0.84, 19.95)$ 和 $(6.57, 8.89, 22.31)$ 进行试验,分别画出第一猜值范围在 $X_0 - 2 < X'_0 < X_0 + 2, Y_0 - 2 < Y'_0 < Y_0 + 2, Z'_0 = Z_0$ 时目标泛函等值线形状图(图 3, 图 4), 并和图 1d, 图 2d 进行比较。对于等值线形状最为复杂的点 $(-0.17, 0.84, 19.95)$ 和点 $(-5.92, -5.90, 24.0)$, 反演难度最大。对点 $(6.57, 8.89, 22.31)$ 的反演难度在某种程度上比以上两点稍好些。最好的是对点 $(-9.42, 9.34, 28.31)$ 的反演,目标泛函形状呈椭圆形。这些试验正反映了 Lorenz 系统的一个特征,即相空间中不同位置对初值敏感程度的不同。下面先给出两个求解个例。

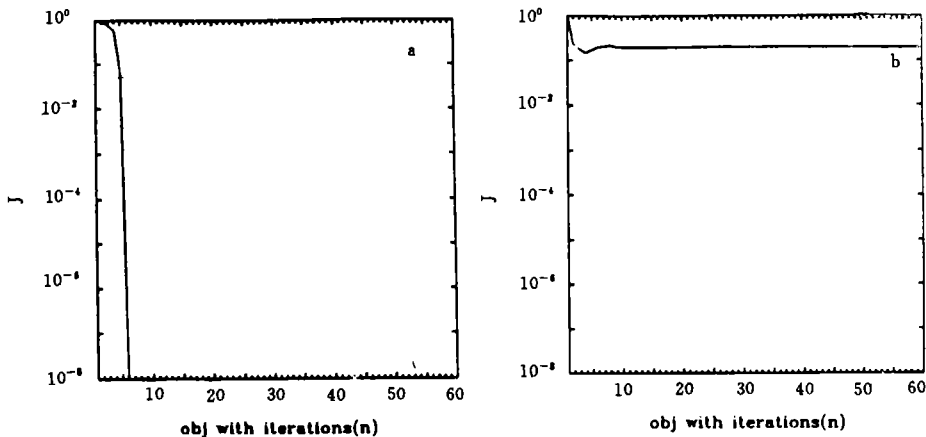


图 5 相空间的两个不同位置目标泛函值随迭代次数的变化

(a 点 $(-9.42, -9.34, 28.31)$; b 点 $(5.92, -5.90, 24.0)$)

4.3 参数 $r=28$ (混沌态)时的两个求解个例

我们给出在 $r=28$ 时,对相空间中两个不同的点为初值的反演情况(实际上当然不

只做了两个,但两个足以说明问题)。第一个例子取的点为 $(X_0, Y_0, Z_0) = (-9.42, -9.34, 28.31)$, 以此为初值积分基本方程, 得到的解作为“观测资料”来做反演试验。第一猜值为 $(-7.54, -7.47, 22.65)$ 时, 图 5a 为反演过程中目标泛函随迭代次数的下降情况(已标准化为 J/J_0), 在 10 步之内下降了 8 个量级。再取不同的第一猜值进行迭代, 试验表明结果相差不多。对此点为初值的反演结果令人满意。第二个例子取准确的初值为 $(X_0, Y_0, Z_0) = (-5.92, -5.90, 24.0)$, 取第一猜值为 $(-10.0, -10.0, 28.0)$, 迭代 60 步, 目标泛函下降不到 1 个量级, 表明结果离真值相差甚远。取不同的第一猜值进行迭代, 其结果均有较大差异, 说明存在有局部极小点, 导致算法失败。

4.4 不精确的 Lorenz 系统的变分同化

由于实际的数值模式都不同程度地存在着误差, 本试验在试验 4.3 节的基础上, 假定 Lorenz 系统的各参数都存在有 10% 的误差, 即 $(\sigma, b, r) = (9.0, 2.4, 25.2)$, 这时 Lorenz 系统的三个方程都是不准确的。控制试验仍以 $(X_0, Y_0, Z_0) = (-9.42, -9.34, 28.31)$ 为初值积分精确的 Lorenz 方程, 得到的解作为“观测资料”来做反演试验。第一猜值仍为 $(-7.54, -7.47, 22.65)$ 时, 图 6a 为反演前 X 分量在同化时段和预报时段的变化。可见由于同化前初值误差较大, 导致两曲线的演变行为相差甚远。图 6b 只反演初值, 由于没有误差的“观测”的作用, 反演的初值和真值有些接近, 但由于 Lorenz 方程的不准确, 反演前后同化时段尤其是预报时段 X 分量的演变曲线也是相差甚远, 这时的同化作用不大; 如果减小模式误差至 5%, 即 $(\sigma, b, r) = (9.5, 2.5334, 26.6)$, 这时, 同化时段 X 分量的演变曲线和真值曲线相差较小, 但预报时段两者曲线却相差甚大(图略)。

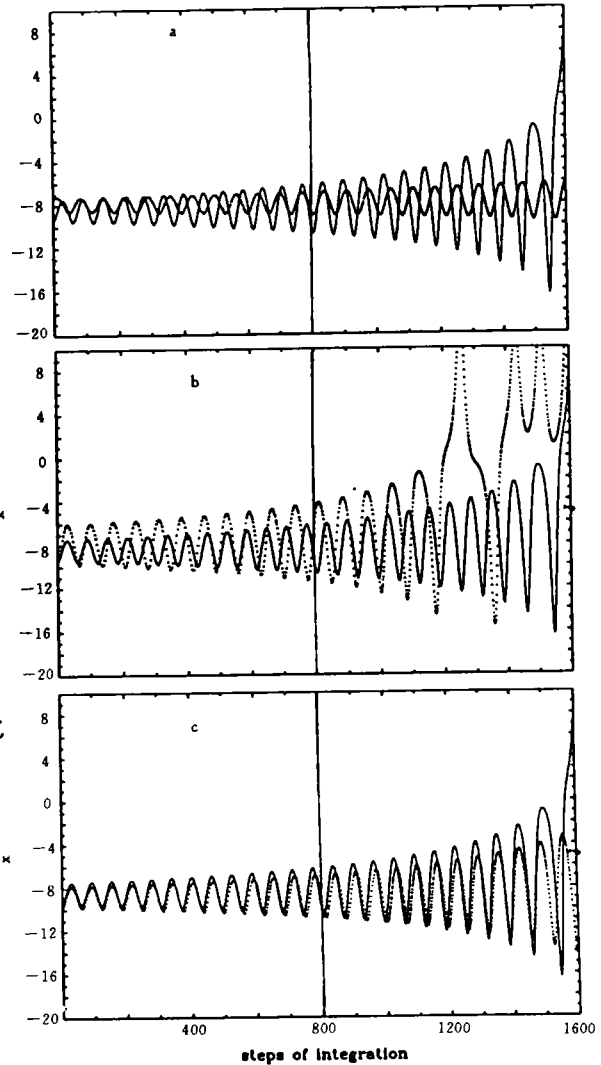


图 6 反演前后同化时段(前 800 步)和预报时段(后 800 步) X 分量的变化曲线

(细实线为由控制试验给出的演变曲线, 图 a 中的粗实线为由第一猜值为初值给出的演变曲线, 图 b 中的虚线为只反演初值后给出的演变曲线, 图 c 中虚线为同时反演初值和参数后给出的演变曲线)

这说明随着方程不精确程度的增加,同化效果愈变愈差。图 6c 我们同时把模式中的参数也作为反演对象,则不仅在同化时段,而且在预报时段内相当长的一段时间内,反演后 X 分量的两曲线都比较接近。这表明同化已使得基于 Lorenz 系统所做预报效果大大提高。附表给出的结果也表明了同化结果和模式参数得到同时改进。

附表 Lorenz 系统初值及各参数的真值第一猜值和反演值的对比

	X	Y	Z	σ	b	r
真值	-9.42	-9.34	28.33	10.0	2.667	28.0
第一猜值	-7.54	-7.47	22.65	9.0	2.4	25.2
反演值	-9.44	-9.37	28.33	9.01	2.698	27.95

4.5 由数值试验得到的启示

至此,我们看到对观测资料的四维同化(即对初值的反演)并不是任何时候都能得到一个和模式初值相协调的初始场。特别是对于存在混沌解的或有较大误差的数值模式,有时能得到和模式相协调的初始场,有时则不能。因此在进行四维同化时,最好也进行下面的工作。

要对模式的动力学性质作些分析,如果模式解对初值非常敏感,模式只有一定的可预报期限,则资料同化的时段必须选取在可预报期限之内。从而可预报期限也可称之为可反演时段。一旦超出此区限,则四维同化的工作可能没有意义。但对于 Lorenz 系统,相空间不同位置,同化效果不同也给我们有益的启示。例如对于中长期数值延伸预报,事先作好可预报技巧的预报对于四维同化也有指导意义。如果本来可预报性太差,则无论用多么好的四维同化方法都难以得到一个和模式相协调的初始场。

实际的数值模式必然存在着误差。如果误差足够大,有可能使同化后得到的解和模式方程不相匹配。在这种情况下,最好是认为模式方程中某些系数也未知,或已知这些系数的初估值,有待进一步修正。同化时对模式初值或模式中系数同时进行反演,则可使同化效果大大改进,这对实际预报模式是具有参考意义的,这也是共轭变分同化的优点之一。

5 结语

本文用 Lorenz 系统作了变分同化的数值试验。即把模式初值作为控制变量来极小化度量了模式解和观测之间距离的目标泛函数,反演最优的初值。通过试验可得到如下一些结论:

(1) 数值模式的非线性作用愈强,模式对初值的敏感性程度愈强,则变分同化的困难愈大。对处于混沌态的 Lorenz 系统,相空间中不同的位置对初值敏感性程度不同,导致可反演性的程度也不同。

(2) 变分同化中的可反演性的概念和数值预报中可预报期限的概念是联系在一起的,可预报性好的模式,对初值的可反演性也愈强。即比较容易从分布于一段时间的观测资料中尽可能多地提取时间演变的信息,得到和动力学模式相协调的初始场。

(3) 在做变分同化的同时对模式方程一些未知项(如未知参数、强迫项等)也同时进行修正,则可使预报效果大大提高。

应该指出的是,本文所有的结果都是基于 Lorenz 系统的数值试验得出的。对于实际的预报模式,估计结论应该不会相差太大,我们将用实际的大气预报模式做进一步的数值试验。

参考文献

- [1] Lewis J M and Derber J C. The use of adjoint equation to solve a variational adjustment problem with advective constraints. *Tellus Ser A*. 1985, 37: 309—322.
- [2] Talagrand O and Courtier P. Variational assimilation of meteorological observation with the adjoint vorticity equation I: theory. *Q J R Met Soc*. 1987, 113: 1311—1328.
- [3] Pierre G. Chaos and quadri-dimensional data assimilation: a study base on Lorenz model. *Tellus Ser A*. 1992, 44: 2—17.
- [4] Stensrud D J and Bao J W. Behaviour of variational and nudging assimilation techniques with a chaotic low-order model. *Mon Wea Rev*. 1992, 120: 3016—3028.
- [5] 罗乔林, 纪立人, 朱宗申. 四维同化的一些方法. 全国第一次资料同化研讨会(摘要).
- [6] 郑维行, 王声望. 泛函分析概要. 高等教育出版社. 1985.
- [7] Navon I M and Legler D. Conjugate-gradient methods for large-scale minimization in meteorology. *Mon Wea Rev*. 1987, 115: 1429—1502.
- [8] 部吉东, 丑纪范. 数值天气预报的两类反问题及一种数值解法—理想试验. *气象学报*. 1994, 52 (2) : 129—137.

THE EFFECTS OF THE MODEL SENSITIVITY TO INITIAL CONDITION UPON THE VARIATIONAL FOUR-DIMENSIONAL ASSIMILATION —— THE STUDY BASED ON LORENZ MODEL

Gao Jidong Chou Jifan

(The Atmospheric Department of Lanzhou University, Lanzhou, 730000)

Abstract

In this paper, some numerical experiments of adjoint variational assimilation has been performed using the famous Lorenz model. With the increase of model sensitivity to initial condition, finding the initial values which harmonize with the Lorenz model become more and more difficult until the scheme is failure on some situation. This shows that the four-dimensional assimilation have some relation with predictability. On the other hand, with the increase of errors in the Lorenz model, the effects of variational assimilation become worse and worse until the forecasting has no meaning. If we perform variational assimilation and retrieve the model parameters at the same time. The consequene of forecasting can improve greatly based on Lorenz model.

Key words: Variational assimilation, Lorenz model, Adjoint equation.