数据同化中的伴随方法及其有关问题的研究*

吕咸青^{1,2} 刘长剑^{2,3}

(¹ 中国科学院海洋研究所 青岛 266071)

(² 物理海洋教育部重点实验室 青岛 266003)

(³ 青岛海洋大学海洋环境学院 266003)

提要 论述了数据同化中的伴随方法与泛函分析中的伴随算子的关系,指出经典的拉格朗日乘法可以作为伴随方法的理论基础。关于伴随方法应用中只能利用模式的伴随的观点,作者认为是有疑问的,并提出了新的观点。所作的数值模拟实验表明,方程的伴随有时明显优于模式的伴随,因而说明方程的伴随也应该引起足够的重视。

关键词 数据同化,变分分析,伴随方法

在人们研究大气、海洋中的有关自然现象的物理机制的过程中,观测资料的充分及合理利用无疑起着至关重要的作用。近些年来,人们发展了一些数据同化的方法,如多项式内插方法,统计最优插值方法以及变分分析方法等。而变分分析方法中的伴随方法是近十余年来发展起来并得到了广泛应用的数据同化方法^[1]。本文将对伴随方法的有关问题作一些探讨。

伴随方法是变分原理和最优控制论相结合的一种方法。伴随方法在海洋研究中的应用是以海洋数学模型作为约束条件,由可观测到的海洋要素反演无法观测到的海洋要素,由海洋的表面信息推断海洋的内部结构。近十余年来的大量研究成果表明伴随方法在数据同化中占有越来越重要的地位^[1]。伴随方法可用来解决许多不同类型的实际问题,其中包括:对模型中的参数进行优化、对初始条件或边界条件进行优化、对定常流或环流问题以及表面热通量问题进行研究等^[2]。

然而,尽管伴随方法应用于研究大气和海洋已有十余年的历史,并能够有效地解决许多类实际问题,但在伴随方法的理论及应用方面仍存在一些疑问,主要表现在以下两个方面:其一是伴随方法究竟与泛函分析中的伴随算子有什么关系?其二是模式的伴随是否优于方程的伴随?第一个问题与伴随方法的理论基础有关,第二个问题与伴随方法的应用有着直接的关系。Thacker 在文献中曾涉及到这两个问题,相关的论据为:其一是强调经典的拉格朗日

乘法简化了伴随算子的推导;其二是强调数据同化只能利用模式的伴随。作者认为 Thacker 对第一个问题有所涉及但并没有给出完整的回答,而他们对第二个问题的回答是有疑问的。以下是对这两个问题的探讨。

1 数据同化中的伴随方法

1.1 泛函分析中伴随算子的定义

A 是 Hilbert 空间 H 上的线性算子,在 H 上定义内积 $\langle \cdot, \cdot \rangle$; 则称 A 为伴随算子,如果 $\langle Ax, y \rangle = \langle x, A^*y \rangle, \forall x, y \in H$ 。我们可以看到泛函分析中只有线性算子才有其伴随算子,而对于非线性算子并无伴随算子之说。

1.2 伴随方法的描述

伴随方法是建立在严格的数学基础之上的一种方法,它将所要解决的实际问题作为条件最小值问题来解决。在此过程中流体力学方程组及其初值、边值条件作为约束条件,使得根据具体问题而设计的代价函数达到最小。

设依赖于时间变量的海洋数学模型为

$$\frac{\partial x}{\partial t} = F(x, c) \quad (1)$$

* 国家重点基础研究发展规划 G1999043808 和国家海洋 863 计划(818-06-05-02)资助项目;中国科学院海洋研究所调查研究报告第 4086 号。

收稿日期:2000-02-15;修回日期:2000-05-22

这里向量 x 的分量是随时间变化的量(如速度、高度、温度、盐度等), t 表示时间, 向量 c 表示模型中的参数, F 是线性或非线性算子。模型方程(1)不封闭, 还需加上相应的定解条件(初始条件、边界条件等)才能确定其唯一解。记 y 表示控制变量, 它包括初始条件、边界条件、模型中的参数等。 y 一旦确定模型(1)的唯一解 $x(y)$ 也就随之确定。当然控制变量必须属于一个可容许的控制集合 Y_{ad} 。 Y_{ad} 可根据有关控制变量的物理学性质及历史经验来确定。

我们的兴趣在于得到模型方程(1)的接近观测结果 x' 的解。这里的所谓接近用代价函数 J 度量模型方程(1)的解与观测结果之间的距离来定义。观测结果包括速度、高度、温度、盐度等。这样, 变分问题可简述为问题(P): 寻求属于 Y_{ad} 的 y^* 使得代价函数 J 达到最小, y^* 表示最优的 y 。

以上所述问题是以模型方程(1)作为约束的约束最小值问题。此问题可转化为无约束最小值问题来求解。当把约束条件作为强约束条件来对待时, 经典的拉格朗日乘子法为我们提供了坚实的理论基础。可构造拉格朗日函数 $L(x, \lambda, y)$

$$L(x, \lambda, y) = J(x, y) + \langle \lambda, G(x, c) \rangle \quad (2)$$

这里 $G(x, c) = \frac{\partial x}{\partial t} - F(x, c)$, λ 的分量是拉格朗日乘子, $\langle \cdot, \cdot \rangle$ 是定义在 $G(x, c)$ 所属 Hilbert 空间上的内积。这样约束最小值问题(P)就被转化成关于 x, λ 和 y 的无约束问题, 而且观测结果将尽可能地接近约束条件的精确解。在此需说明的是这里的约束条件并不是固定不变的, 而是在同化过程中不断得到调整, 使得观测结果尽可能地接近约束条件的精确解。在强约束条件 $G(x, c) = 0$ 的限制下, 确定泛函 $J(x, y)$ 的驻点与确定拉格朗日函数关于变量 x, λ 和 y 的驻点是等价的。表示拉格朗日函数的驻点的方程组称为约束最小值问题(P)的欧拉-拉格朗日方程组。欧拉-拉格朗日最优条件(最优的 x^*, λ^*, y^*)将由下述方程组确定:

$$\frac{\partial L}{\partial \lambda}(x^*, \lambda^*, y^*) = 0 \quad (3)$$

$$\frac{\partial L}{\partial x}(x^*, \lambda^*, y^*) = 0 \quad (4)$$

$$\frac{\partial L}{\partial y}(x^*, \lambda^*, y^*) = 0 \quad (5)$$

方程(3)就是原来的模型方程。方程(4)中的算子是方程(3)中的算子的伴随算子, 通常称方程(4)

为方程(3)的伴随方程。模型方程随时间正向传播信息, 而伴随方程随时间反向传播信息。至于方程(5), 给出了拉格朗日函数关于控制变量的梯度。从以上方程可以看出整个的变分分析过程中, 泛函 J 和约束 $G=0$ 的选取是非常重要的。

1.3 伴随方法与泛函分析中的伴随算子的关系

从以上论述中可以看出, 伴随方法所指的伴随在本质上是经典拉格朗日乘子概念的推广。尽管在某些特殊情形(线性), 泛函分析中的伴随算子可以用来导出伴随方程, 但对于非线性算子而言, 泛函分析中的伴随算子概念就没有任何的可借鉴之处了。

1.4 代价函数 J

数据同化的目的在于寻求模型方程的与观测结果接近的解。我们定义代价函数来度量模型的解与观测结果的距离(差异), 因此要求它具备范数的性质。由于伴随方法中要用到代价函数的导数, 因此又要求它必须是可微分的。通常选择 J 为范数的平方, 将其定义为:

$$J(x, c) = \frac{1}{2}(x - x')^T K_x(x - x') + \frac{1}{2}(c - c')^T K_c(c - c') \quad (6)$$

这里 x' 和 c' 分别表示观测值和估计值, K_x 和 K_c 分别是 x 和 c 所对应的权重矩阵(正定对称矩阵)。(6)中的第一项称谓数据差异, 它将作为驱动伴随方程的外力。(6)式中的第二项之所以作为代价函数的一部分是因为同时也要对模型中的参数进行估计。通过对代价函数的最小化, 既可得到接近观测结果的解, 又可同时优化模型中的参数。换句话说, 对代价函数的最小化过程是通过优化模型中的参数而实现模型方程的解与观测结果的充分接近。

1.5 关于模式的伴随与方程的伴随

尽管伴随方法是建立在严格的数学基础之上而且在理论上并不复杂, 但是不难看出, 利用伴随算子法解决实际问题最难的一点在于: 在每一步同化当中须求解一个正问题, 然后再反向求解伴随方程(计算量相当于正问题的求解)。如何减少计算所需时间以及能否减少计算时间成为数据同化能否取得成功的关键。在利用伴随算子法解决实际问题的过程中, 极其重要的环节是如何得到恰当的伴随方程以及有关控制参数的方程。有两种途径解决这个问题: 一个途径是利用原方程的连续形式去导出其对应的伴随方程, 并分别去匹配各自相应的差分格式

(我们将其称为方程的伴随);另一个途径是利用差分格式去导出相应的伴随差分格式(我们将其称为模式的伴随)。从实际的操作过程来看,两种途径分别导出的伴随差分格式及其参数校正关系是不同的;但究竟哪个好些在理论上是极其复杂的。曾经有学者强调数据同化只能利用模式的伴随,这对于近十年来的伴随方法的运用产生了重要的影响。但我们认为他们的观点是值得商榷的:首先他们所阐述的理由并不充分,其次作者所作的数值模拟实验表明有时方程的伴随明显优于模式的伴随。

作者利用伴随方法作了数值实验:利用“观测结果”来反演海洋涡动黏性系数(10层)和表面风应力系数。尝试了以上所说的两种途径。结果都能够将涡动黏性系数和风应力系数成功地反演出来。所谓的成功指的是:代价函数、数据差异、梯度模几乎都能达到起步值的百分之一、千分之一甚至万分之一以下。在此需要说明的是曾经有人作过这个数值实验,以下的数值实验中的参数与其取了相同的值,但我们得到了更好的实验结果。

2 数值实验

2.1 模型描述

利用修正的 Ekman 模型来反演海洋涡动黏性系数(10层)和表面风应力系数。修正的 Ekman 模型所对应的方程是:

$$\begin{aligned} \frac{\partial u}{\partial t} - fv &= \frac{\partial}{\partial z} \left(A \frac{\partial u}{\partial z} \right) \\ \frac{\partial v}{\partial t} + fu &= \frac{\partial}{\partial z} \left(A \frac{\partial v}{\partial z} \right) \end{aligned} \quad (7.1)$$

相应的边界条件为

$$\begin{aligned} \frac{\partial u}{\partial z} \Big|_{z=0} &= \frac{C_D}{A} \cdot \frac{\rho_a}{\rho_w} \sqrt{u_a^2 + v_a^2} u_a, \\ \frac{\partial v}{\partial z} \Big|_{z=0} &= \frac{C_D}{A} \cdot \frac{\rho_a}{\rho_w} \sqrt{u_a^2 + v_a^2} v_a \end{aligned} \quad (7.2)$$

$$A \frac{\partial u}{\partial z} \Big|_{z=-H_0} = 0, \quad A \frac{\partial v}{\partial z} \Big|_{z=-H_0} = 0 \quad (7.3)$$

初始条件为

$$u \Big|_{t=0} = u_0, \quad v \Big|_{t=0} = v_0 \quad (7.4)$$

上述各式中 H_0 表示 Ekman 层的深度, z 轴垂直向上为正向并且海洋表面处为 $z=0$, f 表示 Coriolis 参数(取为常值), $u(z)$ (向东为正向)和 $v(z)$ (向北为正向)表示海洋中各处的水平方向的运动速

度的分量, $A(z)$ 表示海洋涡动黏性系数; ρ_w 表示海水的密度, ρ_a 表示空气的密度, u_a 和 v_a 表示风速的分量, C_D 表示拖曳系数。

通过变量替换

$$t' = \frac{t}{T_f}, \quad u' = \frac{u}{U}, \quad v' = \frac{v}{V}, \quad z' = \frac{z}{D}$$

$$A' = \frac{A}{S_a}, \quad C'_D = \frac{C_D}{S_c}, \quad u'_a = \frac{u_a}{U_a}, \quad v'_a = \frac{v_a}{U_a}$$

$$\text{这里 } T_f = \frac{1}{f}, \quad D = \left(\frac{S_a}{f}\right)^{1/2}, \quad U = \frac{\rho_a}{\rho_w} \frac{S_c U_a^2}{\sqrt{S_a} f}$$

得到方程

$$\begin{aligned} \frac{\partial u}{\partial t'} - v' &= \frac{\partial}{\partial z'} \left(A' \frac{\partial u}{\partial z'} \right), \\ \frac{\partial v}{\partial t'} + u' &= \frac{\partial}{\partial z'} \left(A' \frac{\partial v}{\partial z'} \right) \end{aligned} \quad (8.1)$$

及边界条件

$$\begin{aligned} A' \frac{\partial u}{\partial z'} \Big|_{z'=0} &= C'_D \sqrt{u_a'^2 + v_a'^2} u_a', \\ A' \frac{\partial v}{\partial z'} \Big|_{z'=0} &= C'_D \sqrt{u_a'^2 + v_a'^2} v_a' \end{aligned} \quad (8.2)$$

$$A' \frac{\partial u}{\partial z'} \Big|_{z'=-H} = 0, \quad A' \frac{\partial v}{\partial z'} \Big|_{z'=-H} = 0, \quad (8.3)$$

和初始条件

$$u \Big|_{t'=0} = u_0, \quad v \Big|_{t'=0} = v_0 \quad (8.4)$$

2.2 连续变分分析:方程的伴随

构造代价函数

$$\begin{aligned} J(u, v, A, C_D) &= \frac{1}{2} K_m \int_t \int_z ((u - \hat{u})^2 \\ &\quad + (v - \hat{v})^2) d\zeta d\tau \\ &\quad + \frac{1}{2} K_a T \int_z (A - \hat{A}) d\zeta \\ &\quad + \frac{1}{2} K_c TH (C_D - \hat{C}_D)^2 \end{aligned} \quad (9)$$

这里 $\hat{u}$ 和 $\hat{v}$ 表示观测值, $\hat{A}$ 和 $\hat{C}_D$ 表示估计值; K_m , K_a 和 K_c 表示权重系数, T 表示整个积分时间。代价函数由 3 项组成,其中第一项为数据差异。

构造 Lagrange 函数

$$L(u, v, A, C_D, \lambda, \mu) = J + \int_t \int_z [\lambda \left(\frac{\partial u}{\partial t} - v - \frac{\partial}{\partial z} \left(A \frac{\partial u}{\partial z} \right) \right) + \mu \left(\frac{\partial v}{\partial t} + u - \frac{\partial}{\partial z} \left(A \frac{\partial v}{\partial z} \right) \right)] d\zeta d\tau \quad (10)$$

我们的目的是寻求在满足方程组(8.1)及其初值、边值条件的限制下,代价函数达到最小。为求得最小值,要求

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial \lambda} = 0,$$

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial \mu} = 0 \quad (11)$$

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial u} = 0,$$

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial v} = 0 \quad (12)$$

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial A} = 0,$$

$$\frac{\partial L(u, v, A, C_D, \lambda, \mu)}{\partial C_D} = 0 \quad (13)$$

方程组(11)对应方程组(8.1),而方程组(12)导出伴随方程,方程组(13)导出关于控制参数 A 和 C_D 的方程。

伴随方程

$$\begin{aligned} \frac{\partial \lambda}{\partial t} - \mu + \frac{\partial}{\partial z} (A \frac{\partial \lambda}{\partial z}) &= K_m(u - \hat{u}), \\ \frac{\partial \mu}{\partial t} + \lambda + \frac{\partial}{\partial z} (A \frac{\partial \mu}{\partial z}) &= K_m(v - \hat{v}), \end{aligned} \quad (14.1)$$

边界条件

$$\frac{\partial \lambda}{\partial z} \Big|_{z=0} = 0, \quad \frac{\partial \mu}{\partial z} \Big|_{z=0} = 0 \quad (14.2)$$

$$\frac{\partial \lambda}{\partial z} \Big|_{z=-H} = 0, \quad \frac{\partial \mu}{\partial z} \Big|_{z=-H} = 0 \quad (14.3)$$

初始条件

$$\lambda|_{t=T} = 0, \quad \mu|_{t=T} = 0 \quad (14.4)$$

控制参数:涡动黏性系数的方程

$$\frac{\partial J}{\partial A} + \int_t \int_z (\frac{\partial \lambda \partial u}{\partial z \partial z} + \frac{\partial \mu \partial v}{\partial z \partial z}) d \xi d \tau = 0 \quad (15)$$

控制参数:风应力系数的方程

$$\begin{aligned} \frac{\partial J}{\partial C_D} - \int_t \int_z (\sqrt{\mu_a^2 + v_a^2} (\lambda|_{z=0} u_a + \mu|_{z=0} v_a)) d \tau \\ = 0 \end{aligned} \quad (16)$$

代价函数的梯度模

$$\Delta J = \sqrt{(\frac{\partial J}{\partial A})^2 + (\frac{\partial J}{\partial C_D})^2} \quad (17)$$

控制参数的校正:涡动黏性系数

$$A_j = \tilde{A}_j + \frac{1}{TK_a} \frac{\partial J}{\partial A_j}, j=1 \dots, 10 \quad (18)$$

控制参数的校正:风应力系数

$$C_D = \tilde{C}_D + \frac{1}{THK_C} \frac{\partial J}{\partial C_D} \quad (19)$$

2.3 有限差分格式

取与微分方程(8.1)对应的有限差分格式 (Crank- Nicholson)

$$\frac{u_j^{i+1} - u_j^i}{\Delta t} - \frac{v_j^{i+1} + v_j^i}{2} = \frac{A_j + A_{j+1}}{4 \Delta z} \frac{u_{j+1}^{i+1} - u_j^{i+1}}{\Delta z}$$

$$- \frac{A_{j-1} + A_j}{4 \Delta z} \frac{u_j^{i+1} - u_{j-1}^{i+1}}{\Delta z}$$

$$+ \frac{A_j + A_{j+1}}{4 \Delta z} \frac{u_{j+1}^i - u_j^i}{\Delta z}$$

$$- \frac{A_{j-1} + A_j}{4 \Delta z} \frac{u_j^i - u_{j-1}^i}{\Delta z},$$

$$\frac{v_j^{i+1} - v_j^i}{\Delta t} - \frac{u_j^{i+1} + u_j^i}{2} = \frac{A_j + A_{j+1}}{4 \Delta z} \frac{v_{j+1}^{i+1} - v_j^{i+1}}{\Delta z}$$

$$- \frac{A_{j-1} + A_j}{4 \Delta z} \frac{v_j^{i+1} - v_{j-1}^{i+1}}{\Delta z}$$

$$+ \frac{A_j + A_{j+1}}{4 \Delta z} \frac{v_{j+1}^i - v_j^i}{\Delta z}$$

$$- \frac{A_{j-1} + A_j}{4 \Delta z} \frac{v_j^i - v_{j-1}^i}{\Delta z} \quad (20)$$

结合边界条件(8.2)和(8.3),我们可得到求解初边值问题(8)的递推关系式

$$MW_i + M_0 W_{i-1} - F_i = 0 \quad (21)$$

这里

$$W = (u_1, u_2, \dots, u_{10}; v_1, v_2, \dots, v_{10})^T,$$

$$F_i = (f_1(i), 0, \dots, 0; f_{11}(i), 0, \dots, 0)^T,$$

$$f_i(i) = \frac{C_D}{A_1} \Delta z \sqrt{\mu_a^2(i \Delta t) + v_a^2(i \Delta t)} u_a(i \Delta t, \Delta t),$$

$$f_{11}(i) = \frac{C_D}{A_1} \Delta z \sqrt{\mu_a^2(i \Delta t) + v_a^2(i \Delta t)} v_a(i \Delta t, \Delta t) \quad (22)$$

矩阵 M 的非零元素为

$$a_{1,1} = 1, a_{1,2} = -1,$$

$$a_{j,j-1} = 0.225(A_{j-1} + A_j),$$

$$a_{j,j} = -1 - 0.225(A_{j-1} + 2A_j + A_{j+1}),$$

$$a_{j,j+1} = 0.225(A_j + A_{j+1}),$$

$$a_{j,j+10} = 0.5 \Delta t, j=2, \dots, 9,$$

$$a_{10,9} = -1, a_{10,10} = 1; a_{11,11} = 1, a_{11,12} = -1,$$

(23)

$$a_{j,j-1} = 0.225(A_{j-11} + A_{j-10}),$$

$$a_{j,j} = -1 - 0.225(A_{j-11} + 2A_{j-10} + A_{j-9}),$$

$$a_{j,j+1} = 0.225(A_{j-10} + A_{j-9}),$$

$$a_{j,j-10} = -0.5 \Delta t, j=12, \dots, 19,$$

$$a_{20,19} = -1, a_{20,20} = 1$$

矩阵 M_0 的非零元素为

$$a_{j,j-1} = 0.225(A_{j-1} + A_j),$$

$$a_{j,j} = 1 - 0.225(A_{j-1} + 2A_j + A_{j+1}),$$

$$a_{j,j+1} = 0.225(A_j + A_{j+1}), a_{j,j+10} = 0.5 \Delta t, j = 2, \dots, 9,$$

$$\begin{aligned} a_{j,j-1} &= 0.225(A_{j-11} + A_{j-10}), \\ a_{j,j} &= 1 - 0.225(A_{j-11} + 2A_{j-10} + A_{j-9}), \\ a_{j,j+1} &= 0.225(A_{j-10} + A_{j-9}), \\ a_{j,j-10} &= -0.5 \Delta t, j = 12, \dots, 19 \end{aligned} \quad (24)$$

与(21)式完全类似,我们可得到求解初边值问题(14)的递推关系式

$$M^a W_{\bar{k}} + M_0^a W_{\bar{k}+1} - F_i^a = 0 \quad (25)$$

这里

$$\begin{aligned} W_{\lambda} &= (\lambda_1, \lambda_2, \dots, \lambda_0, \mu_1, \mu_2, \dots, \mu_0)^T \\ F_i^a &= (f_1^a(i), f_2^a(i), \dots, f_{10}^a(i); \\ & f_{11}^a(i), f_{12}^a(i), \dots, f_{20}^a(i))^T, \\ f_1^a(i) &= f_{10}^a(i) = 0, \\ f_j^a(i) &= \Delta t K_m (u(i, j) - \hat{u}(i, j)), j = 2, \dots, 9, \\ f_{11}^a(i) &= f_{20}^a(i) = 0, \\ f_{j+10}^a(i) &= \Delta t K_m (v(i, j) - \hat{v}(i, j)), \\ & j = 12, \dots, 19 \end{aligned} \quad (26)$$

矩阵 M^a 的非零元素为

$$\begin{aligned} a_{1,1} &= 1, a_{1,2} = -1, \\ a_{j,j-1} &= 0.225(A_{j-1} + A_j), \\ a_{j,j} &= -1 - 0.225(A_{j-1} + 2A_j + A_{j+1}), \\ a_{j,j+1} &= 0.225(A_j + A_{j+1}), \\ a_{j,j+10} &= -0.5 \Delta t, j = 2, \dots, 9, \\ a_{10,9} &= -1, a_{10,10} = 1, \\ a_{11,11} &= 1, a_{11,12} = -1, \\ a_{j,j-1} &= 0.225(A_{j-11} + A_{j-10}), \\ a_{j,j} &= -1 - 0.225(A_{j-11} + 2A_{j-10} + A_{j-9}), \\ a_{j,j+1} &= 0.225(A_{j-10} + A_{j-9}), \\ a_{j,j-10} &= 0.5 \Delta t, j = 12, \dots, 19, \\ a_{20,19} &= -1, a_{20,20} = 1, \end{aligned} \quad (27)$$

矩阵 M_0^a 的非零元素为

$$\begin{aligned} a_{j,j-1} &= 0.225(A_{j-1} + A_j), \\ a_{j,j} &= 1 - 0.225(A_{j-1} + 2A_j + A_{j+1}), \\ a_{j,j+1} &= 0.225(A_j + A_{j+1}), \\ a_{j,j+10} &= -0.5 \Delta t, j = 2, \dots, 9; \\ a_{j,j-1} &= 0.225(A_{j-11} + A_{j-10}), \\ a_{j,j} &= 1 - 0.225(A_{j-11} + 2A_{j-10} + A_{j-9}), \\ a_{j,j+1} &= 0.225(A_{j-10} + A_{j-9}), \end{aligned}$$

$$a_{j,j-10} = 0.5 \Delta t, j = 12, \dots, 19 \quad (28)$$

2.4 离散变分分析:模式的伴随

数据差异的离散形式

$$DM = 0.5 K_m \Delta t \Delta z \sum_{i=1}^{480} \sum_{j=1}^{10} [(u_j^i - \hat{u}_j^i)^2 + (v_j^i - \hat{v}_j^i)^2] \quad (29)$$

代价函数的离散形式

$$\begin{aligned} J(W, A, C_D) &= 0.5 K_m \Delta t \Delta z \sum_{i=1}^{480} (W_i - \hat{W}_i)^T \\ & \times (W_i - \hat{W}_i) + 0.5 K_a T \Delta z \sum_{j=1}^{10} \\ & (A_j - \tilde{A}_j) \\ & + 0.5 K_c TH(C_D - \tilde{C}_D)^2 \end{aligned} \quad (30)$$

构造离散形式的 Lagrange 函数

$$\begin{aligned} L(W, A, C_D, W_{\lambda}) &= J + \Delta t \Delta z \sum_{i=1}^{480} M_{\lambda_1}^T (M W_i \\ & + M_0 W_{i-1} - F_i) \end{aligned} \quad (31)$$

我们的目的是寻求在满足方程组(21)的限制下,代价函数达到最小。为求得最小值,要求

$$\frac{\partial L(W, A, C_D, W_{\lambda})}{\partial W_{\lambda_1}} = 0, i = 1, \dots, 480 \quad (32)$$

$$\frac{\partial L(W, A, C_D, W_{\lambda})}{\partial W_i} = 0, i = 1, \dots, 480 \quad (33)$$

$$\frac{\partial L(W, A, C_D, W_{\lambda})}{\partial A_j} = 0, j = 1, \dots, 10 \quad (34)$$

$$\frac{\partial L(W, A, C_D, W_{\lambda})}{\partial C_D} = 0 \quad (35)$$

方程组(32)对应方程组(21),而方程组(33)导出伴随方程,方程组(34)和(35)导出关于控制参数 A 和 C_D 的方程。

伴随方程

$$M^T W_{\lambda_1} + M_0^T W_{\lambda_{1+1}} + K_m (W_i - \hat{W}_i) = 0 \quad (36)$$

控制参数:涡动系数的方程

$$\begin{aligned} \Delta z K_a T (A_1 - \tilde{A}_1) \\ + \Delta t \Delta z \sum_{i=1}^{480} W_{\lambda_1}^T \left(\frac{\partial M}{\partial A_1} W_i + \frac{\partial M_0}{\partial A_1} W_{i-1} \right) \\ + \Delta t \Delta z^2 \sum_{i=1}^{480} \frac{C_D}{A_1^2} \sqrt{u_a^2 + v_a^2} [u_a(i \Delta t) \omega_{\lambda_1}^j \\ + v_a(i \Delta t) \omega_{\lambda_{1+1}}^j] = 0 \end{aligned} \quad (37)$$

$$\begin{aligned} \Delta z K_a T (A_j - \tilde{A}_j) + \Delta t \Delta z \sum_{i=1}^{480} W_{\lambda_1}^T \left(\frac{\partial M}{\partial A_j} W_i \right. \\ \left. + \frac{\partial M_0}{\partial A_j} W_{i-1} \right) = 0, j = 2, \dots, 10. \end{aligned}$$

控制参数:风应力系数的方程

$$K_c TH(C_D - \tilde{C}_D)$$

$$-\Delta t \Delta z^2 \frac{C_D}{A_1} \sum_{i=1}^{480} \sqrt{u_a^2 + v_a^2} [u_a(i \Delta t) \omega_{A1}^j + v_a(i \Delta t) \omega_{A1}^j] = 0 \quad (38)$$

2.5 实验结果

连续变分分析导出方程的伴随。所谓方程的伴随指的是:利用递推关系式(21)来求解初边值问题(8),利用递推关系式(25)来求解伴随初边值问题(14),控制参数的校正利用式(15)和(16)。

离散变分分析导出模式的伴随。所谓方程的伴随指的是:利用递推关系式(21)来求解初边值问题(8),利用递推关系式(36)来代替伴随初边值问题,控制参数的校正利用式(37)和(38)。

要使代价函数、数据差异、梯度模都达到起步值的百分之一以下,方程伴随只需同化13步,而模式的伴随却需要66步。要使代价函数、数据差异、梯度模都达到起步值的千分之一以下,方程伴随只需同化61步,而模式的伴随却需要454步。详细的结

果见下表。表1中的数字为实现优化目标所需的同化步数(在利用模式的伴随时,同化的步数即使达到100 000步,代价函数依然达不到起步值的万分之一以下)。

表1 达到相同优化目标时,两种途径的比较

Tab.1 While obtaining the same optimization goals, two ways are compared

优化目标	方程的伴随	模式的伴随
数据差异首达起步值的百分之一以下	6	37
数据差异首达起步值的千分之一以下	21	97
数据差异首达起步值的万分之一以下	76	415
代价函数首达起步值的百分之一以下	13	66
代价函数首达起步值的千分之一以下	55	454
代价函数首达起步值的万分之一以下	149	100 000*
梯度模首达起步值的百分之一以下	13	28
梯度模首达起步值的千分之一以下	61	100
梯度模首达起步值的万分之一以下	561	310

表2 同化步数及其达到的目标(方程的伴随)

Tab.2 Steps of assimilation and the goals obtained (finite difference of adjoint)

参数	C _D	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉
目标	1.000 0	.045 0	.037 5	.030 5	.025 0	.020 5	.017 0	.014 0	.011 5	.009 5
6	1.025 3	.047 9	.038 3	.028 6	.021 7	.015 1	.011 3	.009 6	.008 5	.007 4
13	1.012 5	0.474	.039 0	.030 7	.024 7	.018 4	.013 5	.010 7	.008 9	.007 1
21	1.007 8	.046 8	.038 8	.030 9	.025 5	.019 9	.015 0	.011 7	.009 4	.007 0
55	1.005 6	.045 8	.038 1	.030 4	.025 2	.020 7	.016 8	.0136	.010 6	.007 6
61	1.002 3	.045 6	.038 0	.030 3	.025 1	.020 7	.0169	.013 7	.010 7	.007 8
76	1.007 1	.045 6	.038 0	.030 5	.025 1	.020 7	.017 0	.014 0	.011 0	.008 1
149	0.999 7	.044 8	.037 7	.030 6	.025 0	.020 6	.016 9	.014 1	.011 5	.008 9
561	0.994 0	.044 5	.037 6	.030 6	.025 1	.020 5	.017 1	.013 8	.011 6	.009 4

表3 同化步数及其达到的目标(模式的伴随)

Tab.3 Steps of assimilation and the goals obtained (adjoint of finite difference)

参数	C _D	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉
目标	1.000 0	.045 0	.037 5	.030 5	.025 0	.020 5	.017 0	.014 0	.011 5	.009 5
28	1.008 6	.046 6	.033 7	.017 3	.020 5	.014 3	.011 3	.009 8	.009 0	.008 4
37	1.010 2	.045 4	.034 3	.020 5	.022 8	.015 4	.012 1	.010 4	.009 4	.008 5
66	1.006 3	.044 5	.035 5	.025 4	.026 0	.017 4	.013 7	.011 5	.010 1	.008 8
97	1.003 9	.044 7	.036 5	.027 4	.027 1	.018 3	.014 7	.012 3	.010 6	.008 9
100	1.003 3	.044 8	.036 5	.027 5	.027 2	.018 4	.014 8	.012 3	.010 6	.008 9
310	0.099 9	.045 6	.038 3	.029 2	.027 1	.019 4	.016 9	.013 8	.011 4	.009 2
415	0.998 4	.015 5	.038 0	.029 1	.026 9	.019 4	.017 1	.013 9	.011 5	.009 3
454	0.998 1	.045 4	.038 0	.028 2	.026 8	.019 4	.017 2	.013 9	.011 5	.009 3
10 ⁵	0.996 6	.045 0	.038 0	.030 2	.025 6	.020 0	.017 5	.013 3	.012 0	.008 9

3 结论

伴随方法还包括基于最优控制理论原理的伴随方法, Talagrand 等人提出共轭变分同化方法是导出控制方程的切变线性方程, 利用泛函分析伴随算子理论导出切线性方程的共轭方程, 利用最优化理论避开求解 Euler-Lagrange 方程的困难。

拉格朗日乘法导出的伴随在很大程度上是泛函分析中的伴随算子概念的借用。借助于拉格朗日乘法可以导出正确的伴随方程, 数据差异自然地加到了伴随方程上作为驱动。

我们认为过分强调模式的伴随是有疑问的。我们的理由主要表现在: 首先, 他们所阐述的理由并不充分: 其理由之一在于离散的动力学模型作为约束, 之二在于观测结果是离散的。不可否认, 他们给出了解决问题的正确方法, 这种方法深刻地影响了近十多年的同化方法在海洋及大气研究中的应用。但这毕竟只是能够取得成功的方法之一, 过分强调模

式的伴随而排斥其他的方法难以经得起实践的检验。其次, 我们所作的数值模拟实验表明有时方程的伴随明显优于模式的伴随; 第三, 通过方程的伴随我们得到的伴随方程有着明显的物理意义以及合理的初始条件和边界条件, 而通过模式的伴随我们得到的伴随方程的物理意义是不明确的, 这对于边界条件尤其如此。

由于非线性算子的伴随算子的推导存在本质上的困难, 也许这是造成人们过分强调模式的主要伴随的原因。我们已找到了切实可行的办法来导出非线性算子的伴随算子, 并将专门对此进行讨论。究竟模式的伴随与方程的伴随哪个更优, 这在理论上是极其复杂的, 最终只有靠实践的检验了。✎

参考文献

- 1 Navon I. M. *Dynamics of Atmospheres and Oceans*, 1997, 27: 55 ~ 79
- 2 Yu L. and Ó'Brien J. J. *Journal of Physical Oceanography*, 1995, 25: 2 319 ~ 2 343

STUDY ON THE ADJOINT METHOD IN DATA ASSIMILATION AND THE RELATED PROBLEMS

LÜ Xian-qing^{1,2} LIU Wen-jian^{2,3}

(¹ Institute of Oceanology, Chinese Academy of Science, Qingdao, 266071)

(² Key Laboratory of Physical Oceanography, Ministry of Education, Qingdao, 266003)

(³ College of Marine Environmental, Ocean University of Qingdao, 266003)

Received: Feb. 15, 2000

Key Words: Data assimilation, Variational analysis, Adjoint method

Abstract

The relations of the adjoint method in data assimilation and the adjoint operator in functional analysis were studied. It is pointed out that the classical Lagrange multiplier method is one of the theoretical fundament of the adjoint method in data assimilation. The author holds it is not reasonable that one can only use the adjoint of model in data assimilation. The simulated numerical experiment shows that the adjoint of equation is sometimes superior to the adjoint of model. (本文编辑:张培新)