Adjoint Matching Condition for Parameterized Discontinuities—A Derivation Using Lagrangian–form Costfunction

Xu Qin (许 秦)
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, USA
and *Qiu Chongjian* (邱崇践)
Department of Atmospheric Sciences, Lanzhou University, Lanzhou 730001
Received July 10, 1996; revised October 5, 1996

ABSTRACT

The generalized adjoint property and adjoint matching condition for systems that contain discontinuous on / off switches are derived by a perturbation analysis of the Lagranging–form costfunction.

Key words: Discontinuity, Adjoint property, Perturbation analysis

I. INTRODUCTION

Discontinuous on / off switches often occur with parameterized processes in atmospheric models. As shown by Vukicevic and Errico (1993), parameterized on / off switches can produce spikes in the perturbation tendency terms and these spikes cannot be captured by the classic tangent linearization. In terms of generalized function, these spikes manifest delta–functions in the tangent linear equation and their "transposes" are delta–functions in the associated adjoint equation (Xu 1995a, b, hereafter referred to as X95a,b). Since these delta–functions were previously ignored in the conventional applications of the classic adjoint method, the associated discontinuous jumps of the tangent linear variable and adjoint variable at the switches are neglected. Because of this, significant errors could be introduced in the computations of the costfunction gradients.

The discontinuous jumps of the tangent linear variable and adjoint variable can be described by their respective matching conditions. In X95b, the tangent linear matching condition was derired by a perturbation analysis that considers the variation of the switch point; the adjoint matching condition was derived from the generalized adjoint property or, say, the adjoint property between the generalized tangent linear operator and adjoint operator; and the generalized adjoint property was derived from the generalized tangent linear and adjoint equations that contain delta–functions. The purpose of this note is to show that the generalized adjoint property and adjoint matching condition of X96b can be also derived by a perturbation analysis of the Lagrangian–form costfunction. The model equations and costfunction formulations are presented in the next section. The detailed derivation is given in Section 3.

II. MODEL EQUATIONS AND COSTFUNCTION FORMULATIONS

A spatially discretized model used in data assimilation can be written into the following vector form:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x},t) \ , \tag{1}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{2}$$

where $\mathbf{x} = \mathbf{x}(\mathbf{x}_0, t)$ *is the model's state at time* t, $\mathbf{x}_0$ *is the initial state, and* $\mathbf{F}$ *can be a nonlinear* function of $(\mathbf{x},t)$. For simplicity, we assume that during the data assimilation period $[0,T]$ there is only a single switch at $t=\tau$ in association with the following threshold condition:

$$c[\mathbf{x}(t), t] \geqslant 0 (\text{or} < 0) \quad \text{for on state (off state)} \ . \tag{3}$$

At this switch point, $\mathbf{F}$ has a discontinuous jump and the amplitude of the jump is given by

$$\mathbf{G} = \mathbf{F}_2(\mathbf{x}, \tau_+) - \mathbf{F}_1(\mathbf{x}, \tau_-) \ , \tag{4}$$

where $\mathbf{F}_2(\mathbf{x},\tau_+)$ and $\mathbf{F}_1(\mathbf{x},\tau_-)$ are the lefthand limit and righthand limit of $\mathbf{F}$, respectively.

The costfunction can be defined as follows:

$$\mathbf{J}(\mathbf{x}_0) = \int_0^T D[\mathbf{x}(\mathbf{x}_0, t), t] dt \ , \tag{5}$$

where D measures the discrepancy between the model state and observation. The objective is to find the optimal estimate of the initial state $\mathbf{x}_0$ that minimizes $\mathbf{J}$ under the constraints of (1)–(4). The minimization problem can be reformulated into an unconstrained form in association with the following Lagrangian–form costfunction:

$$J_L(\mathbf{x}_0) = \int_0^{\tau_-} \{D(\mathbf{x}, t) - \mathbf{a}^T[\frac{d\mathbf{x}}{dt} - \mathbf{F}_1(\mathbf{x}, t)]\} dt + \int_{\tau_+}^T \{D(\mathbf{x}, t)$$

$$- \mathbf{a}^T[\frac{d\mathbf{x}}{dt} - \mathbf{F}_2(\mathbf{x}, t)]\} dt \ , \tag{6}$$

where the vector Lagrangian multiplier $\mathbf{a}$ defines the adjoint vector, and $(\ )^T$ represents the transpose of $(\ )$.

III. PERTURBATION ANALYSIS AND MATCHING CONDITIONS

For a small perturbation $\delta\mathbf{x}_0$ in the initial state, the leading order variation of J_L can be obtained as follows:

$$\delta J_L = \delta J_1 + \delta J_2 + \delta J_3 + \delta J_4 \ , \tag{7}$$

where

$$\delta J_1 = \int_0^{\tau_-} \{\frac{d\mathbf{a}^T}{dt} + \mathbf{a}^T \frac{\partial \mathbf{F}_1(\mathbf{x}, t)}{\partial \mathbf{x}} + [\frac{\partial D(\mathbf{x}, t)}{\partial \mathbf{x}}]^T\} \delta\mathbf{x} dt + \int_{\tau_+}^T \{\frac{d\mathbf{a}^T}{dt}$$

$$+ \mathbf{a}^T \frac{\partial \mathbf{F}_2(\mathbf{x}, t)}{\partial \mathbf{x}} + [\frac{\partial D(\mathbf{x}, t)}{\partial \mathbf{x}}]^T\} \delta\mathbf{x} dt \ , \tag{8}$$

$$\delta J_2 = \{D(\mathbf{x}, \tau_-) - \mathbf{a}(\tau_-)^T[\frac{d\mathbf{x}(\tau_-)}{dt} - \mathbf{F}_1(\mathbf{x}, \tau_-)]\}\delta\tau - \{D(\mathbf{x}, \tau_+)$$

$$- \mathbf{a}(\tau_+)^T[\frac{d\mathbf{x}(\tau_+)}{dt} - \mathbf{F}_2(\mathbf{x}, \tau_+)]\}\delta\tau \ , \tag{9}$$

$$\delta J_3 = \mathbf{a}(\tau_+)^T \delta\mathbf{x}(\tau_+) - \mathbf{a}(\tau_-)^T \delta\mathbf{x}(\tau_-) \ , \tag{10}$$

$$\delta J_4 = \mathbf{a}(0)^T \delta \mathbf{x}(0) - \mathbf{a}(T)^T \delta \mathbf{x}(T) , \tag{11}$$

and $\delta\tau$ is the variation of the switch time. The adjoint equation requies that the integrand in (8) vanishes over $[0, \tau_-)$ and $(\tau_+, T]$, and this gives $\delta J_1 = 0$. It is also obvious that $\delta J_2 = 0$, because $D(\mathbf{x}, t)$ is continuous at $t = \tau$ and the state vector $\mathbf{x}$ satisfies the model's equation (1) on both (lefthand and righthand) sides of τ. As shown in X95b, the matching conditions for the tangent linear and adjoint vectors should satisfy

$$\mathbf{a}(\tau_-)^T \delta \mathbf{x}(\tau_-) = \mathbf{a}(\tau_+)^T \delta \mathbf{x}(\tau_+) , \tag{12}$$

and this makes $\delta J_3 = 0$. Finally, with $\mathbf{a}(T) = 0$ for the adjoint vector, we have

$$\delta J_L = \mathbf{a}(0)^T \delta \mathbf{x}_0 \qquad \text{or} \qquad \nabla_{\mathbf{x}_0} J_L = \mathbf{a}(0) . \tag{13}$$

This shows that the gradient of J_L with respect to $\mathbf{x}_0$ is given by $\mathbf{a}(0)$–the adjoint vector value at $t = 0$. This vector value can be obtained by backward integration of the adjoint equation with the initial condition $\mathbf{a}(T) = 0$ and a matching condition satisfying (12) at the switch point.

The above analysis shows that (12) is the desired condition to ensure the adjoint formulation of the gradient in (13), although the analysis does not give by itself the specific forms of matching conditions for the tangent linear vector and adjoint vector. As shown in X95b, (12) is the generalized adjoint property at the switch point in association with the generalized tangent linear and adjoint operators, and the generalized adjoint property can be expressed explicitly as a whole (over the entire period of data assimilation) by the resolvents of the generalized tangent linear and adjoint operators. Through (12), the matching condition for the adjoint vector is tied up with the matching condition for the tangent linear variable. Note that (7)–(11) are similar to (13) of Bao and Warner (1993), and these results were previously used by Bao and Kuo (1995), but the generalized adjoint property (12) was unnoticed.

The matching condition for the tangent linear vector can be derived by analyzing the variation of the switch point caused by the concerned perturbation, and the detailed derivations are given in X95b. Here we only need to show the result [see (3.9) of X95b]:

$$\delta \mathbf{x}(\tau_+) = (\mathbf{I} + \mathbf{A}) \delta \mathbf{x}(\tau_-) , \tag{14}$$

where $\mathbf{I}$ is the unite matrix, $\mathbf{A} \equiv \mathbf{G}(\nabla_{\mathbf{x}} c)^T / |dc/dt|_-$ is the jump matrix, and $|dc/dt|_- = |\partial c/\partial t + (d\mathbf{x}/dt)^T \nabla_{\mathbf{x}} c|_- = |\partial c/\partial t + \mathbf{F}_1^T \nabla_{\mathbf{x}} c|_-$ is the absolute value of the total time derivative of c at $t = \tau_-$ (lefthand limit). Substituting (14) into (12) gives the following matching condition for the adjoint vector:

$$\mathbf{a}(\tau_-) = (\mathbf{I} + \mathbf{A}^T) \mathbf{a}(\tau_+) . \tag{15}$$

This matching condition is the same as in (4.4) of X95b but different from (2.2) of Bao and Kuo (1995).

To trigger an on switch, c must increase to the threshold vtlue ($c = 0$) according to (3), so $|dc/dt|_- = |(dc/dt)|_- > 0$ for the on switch. As explained in X96b, the on switch (or off switch) is normally determined by the last exceeded (or first not exceeded) threshold condition. In Bao and Kuo (1995), a vector threshold condition was used and all the component threshold conditions were required to be simultaneously satisfied at the switch point for both the reference and perturbed solutions. This imposed unrealistic constraints on the admissible variations of the tangent linear perturbation. Thus, a vector threshold condition is inconsistent with the situation in an atmospheric model, although the same type of vector equaliy was used as interior-point constraints in (3.5.1) of Bryson and Ho (1975) for engineering applications of optimal control theory. When a switch is triggered simultaneously by several threshold conditions, the situation becomes complicated and requires certain special considerations for the application of the matching conditions (14)-(15). The details are beyond the scope of this note.

This work is supported by the NOAA grant NA37RJ0203 and the Natural Science Foundation of China (49375241).

REFERENCES

Bao, J. W., and Y. -H. Kuo (1995). On / off switches in the adjoint method: Step functions, *Mon. Wea. Rev.*, **123**: 1589-1594.

Bryson, A. E., and Y. -C. Ho (1975), *Applied Optimal Control*, Hemisphere Publishing Corporation, 481 pp.

Vukicevic, T., and R. M., Errico (1993), Linearization and adjoint of parameterized moist diabatic processes, *Tellus*, **45A**, 493-510.

Xu, Q. (1996), Generalized adjoint for physical processes with parameterized discontinuities-Part I: Basic issues and heuristic examples, *J. Atmos. Sci.*, 1123-1142.

Xu, Q. (1996), Generalized adjoint for physical processes with parameterized discontinuities-Part II: Vector Formulations and Matching Conditions, *J. Atmos. Sci.*, 1143-1155.