Final version
E. Ustinov 4/02/03

# On the matrix approach to the adjoint sensitivity analysis: Direct treatment of multiple observables

Eugene A. Ustinov
*Jet Propulsion Laboratory, California Institute of Technology*

Sensitivity analysis of models with multiple observables involves, directly or indirectly, sensitivities of individual observables to individual parameters of the models. The conventional adjoint approach to the evaluation of sensitivities of observables uses the adjoint *vector* solution and yields the sensitivities of a single observable (or single linear combination of observables) to multiple parameters. Recently the author has extended the adjoint approach to the sensitivity analysis was extended to using the adjoint *matrix* solution to yield sensitivities of multiple observables to multiple parameters. In this presentation, we compare the vector and matrix adjoint approaches to the sensitivity analysis as applied to the problem of assimilation of data.

## 1. Matrix approach and vector of observables

## 2. Vector approach and scalar cost function

## 3. Comparison of the matrix and vector approach

## 4. Conclusion

Linearization of Eqs. (5), (6) yields:

$$\mathbf{R}'_e = \int_{t_0}^{t_1} \mathsf{W}_e^{\mathsf{T}}(t)\mathbf{V}'(t)\, dt' = (\mathsf{W}_e,\, \mathbf{V}') \tag{7}$$

where $(\mathsf{W},\, \mathbf{V}')$ is a corresponding inner product, and

$$\mathbf{R}'_c = \mathsf{W}_c^{\mathsf{T}}\, \mathbf{V}'(t_1) \tag{8}$$

As shown in [1], perturbation of observables $\mathbf{R}'$ due to variations of model parameters contained in Eq. (3) can be expressed in the form of an inner product

$$\mathbf{R}'_e = (\mathsf{V}^*,\, \mathbf{S_e}) \tag{9}$$

where, $\mathsf{V}^*$ is the solution of the adjoint problem

$$-\frac{d\mathsf{V}^*}{dt} + \mathsf{C}^{\mathsf{T}}(t)\mathsf{V}^*(t) = \mathsf{W}_e(t) \tag{10}$$

$$\mathsf{V}^*|_{t_1} = \mathsf{W}_c \tag{11}$$

Also, it can be shown that

$$\mathbf{R}'_c = \mathsf{V}^{*\mathsf{T}}(t_0)\, \mathbf{S}_c \tag{12}$$

(NB! In [1] the expression Eq. (9) contains an extra term which, as was overlooked in [1], is quadratic in perturbations of model parameters and can safely be omitted in the linear approximation.) Corresponding sensitivites have the form:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{a}} = \left(\mathsf{V}^*,\, \frac{\partial \mathbf{S_e}}{\partial \mathbf{a}}\right) \tag{13}$$

$$\frac{\partial \mathbf{R}}{\partial \mathbf{S}_c} = \mathsf{V}_c^{*\mathsf{T}}(t_0) \tag{14}$$

## 2. Vector approach and scalar cost function

Now we consider the conventional approach to the adjoint sensitivity analysis as it is used in the variational assimilation of data. It was founded in the pioneering paper by Talagrand and Courtier (1987) [2] and its principles remain unchanged since then (see, e.g., (Talagrand, 1997) [3]). The nonlinear initial value problem here has the form (cf. Eq. (5) of [2]):

$$\frac{d\mathbf{V}}{dt} + \mathcal{N}[\mathbf{V}] = \mathbf{0} \tag{15}$$

$$\mathbf{V}|_{t_0} = \mathbf{V}_0 \tag{16}$$

(All model parameters are limited to the right-hand term of the initial conditions.) Corresponding linearized forward problem has the form (cf. Eqs. (3), (4):

$$\frac{d\mathbf{V}'}{dt} + \mathsf{C}(t)\mathbf{V}'(t) = 0 \tag{17}$$

$$\mathbf{V}'|_{t_0} = \mathbf{S}_c \tag{18}$$

Further on, an artificial scalar observable is introduced in [2], in the form of a squared, not specified, norm $H[\mathbf{V}'(t), t]$ of the perturbation $\mathbf{V}'$

$$R = \int_{t_0}^{t_1} H[\mathbf{V}'(t), t]\, dt \tag{19}$$

For illustration purposes, we also consider the simplified, Euclidean norm:

$$R = \int_{t_0}^{t_1} \mathbf{V}'^{\mathsf{T}}(t)\mathbf{V}'(t)\, dt = (\mathbf{V}', \mathbf{V}') \tag{20}$$

The scalar observable $R$ defined in such a way, is essentially nonlinear and has to be linearized first. Considering the variation

$\delta\mathbf{V}$ of the linearized forward solution $\mathbf{V}'(t)$ we linearize Eqs. (19), (20) in the form:

$$\delta R = \int_{t_0}^{t_1} (\nabla_{\mathbf{V}} H, \delta\mathbf{V}(t))\, dt = \int_{t_0}^{t_1} (\nabla_{\mathbf{V}} H)^{\mathsf{T}} \delta\mathbf{V}(t)\, dt \tag{21}$$

[cf Eq. (7) of (Talagrand and Courtier 1987)] and

$$\delta R = 2 \int_{t_0}^{t_1} \mathbf{V}'^{\mathsf{T}}(t) \delta\mathbf{V}(t)\, dt \tag{22}$$

whence, for the Euclidean norm used here,

$$\nabla_{\mathbf{V}} H = 2\,\mathbf{V}'(t) \tag{23}$$

Comparing Eqs. (7) and (22) we see that the observables weighting function $\mathsf{W}_e(t)$ in (Talagrand and Courtier 1987) is defined as

$$\mathsf{W}_e(t) = (\nabla_{\mathbf{V}} H)^{\mathsf{T}} \tag{24}$$

Or, with the norm Eq.19 defined as in Eq.20, the function $\mathsf{W}_e(t)$ is defined by the linearized solution $\mathbf{V}'(t)$:

$$\mathsf{W}_e(t) = 2\,\mathbf{V}'^{\mathsf{T}}(t) \tag{25}$$

Finally, since we have no observables at $t = t_1$, the matrix $\mathsf{W}_c$ is equal zero and, in accordance with (Talagrand and Courtier 1987), we have the final value condition in the form:

$$\mathsf{W}_c = \mathbf{0} \tag{26}$$

## 3. Comparison of the matrix and vector approach

In this presentation, we argue that the ability to evaluate sensitivities of individual observables to model parameters, which is intrinsic to the matrix approach [1], is important because it facilitates:

(1) Selection of the most informative observables in the data sets and discarding of less informative ones;

(2) Optimization of the return of planned measurements by maximizing of the information content of data to be obtained;

(3) Reduction of the number of iterations in the variational data assimilation process.

Below we consider these benefits separately.

### *3.1. Selection of the most informative observables*

Under the matrix approach, the sensitivities of individual observables to individual model parameters are directly obtained from Eqs. (13), (14). Under the conventional approach, the individual observables are reduced to the vector of differences $\mathbf{V}'$ between observed and modelled values of the state vector $\mathbf{V}$, and, moreover, these differences are combined into one, artificial scalar observable $R$, Eqs. (19), (20).

### *3.2. Optimization of the return of planned measurements*

Equations (19), (20) provide a possibility to pre-compute, forecast sensitivities of results of planned measurements (observables) to the model parameters of interest. Choosing the observables with maximum sensitivities, one can discard the data adding little or nothing to the information content. This promises to reduce the requirements to the computer resources and to decrease the computer time expenses.

The conventional approach does not provide this possibility.

*3.3. Reduction of the number of iterations*

For demonstration purposes we use a simple model where the observables coincide with the elements of the state vector at the end of the integration interval $\mathbf{V}(t_1)$, and the model parameters are limited to the elements of the initial state vector $\mathbf{V}(t_0) = \mathbf{S}_c$.

Following the matrix approach we deal with the vector of linearized observables (cf. Eq. (8)):

$$\mathbf{R}' = \mathsf{I}^{\mathsf{T}} \mathbf{V}'(t_1) \tag{27}$$

where $\mathsf{I}$ is a unix matrix of corresponding rank. Corresponding adjoint problem is:

$$-\frac{d\mathsf{V}^*}{dt} + \mathsf{C}^{\mathsf{T}}(t)\mathsf{V}^*(t) = 0 \tag{28}$$

$$\mathsf{V}^*|_{t_1} = \mathsf{I} \tag{29}$$

and its matrix solution yields the matrix of sensitivities (cf Eq. (14))

$$\frac{\partial \mathbf{R}}{\partial \mathbf{V}(t_0)} = \mathsf{V}^{*\mathsf{T}}(t_0) \tag{30}$$

to be used in the variational data assimilation.

Following the conventional approach, we deal with an artificial scalar observable

$$R' = \mathbf{V}'^{\mathsf{T}}(t_1)\mathbf{V}'(t_1) \tag{31}$$

Corresponding adjoint problem is (cf. Eq. (25)):

$$-\frac{d\mathbf{V}^*}{dt} + \mathsf{C}^{\mathsf{T}}(t)\mathbf{V}^*(t) = 0 \tag{32}$$

$$\mathbf{V}^*|_{t_1} = \mathbf{V}'(t_1) \tag{33}$$

and its vector solution yields the row vector of sensitivities of $R'$ (cf Eq. (14))

$$\frac{\partial R}{\partial \mathbf{V}(t_0)} = \mathbf{V}^{*\mathsf{T}}(t_0) \tag{34}$$

to be used in the variational data assimilation.

Now we compare the procedures of variational assimilation using two approaches. In both cases, the vector $\mathbf{V}'$ is treated as a residual, the difference between the observed value and modelled value of the state vector $\mathbf{V}$. This difference is used to infer the correction to the initial value of the state vector $\mathbf{V}'_0$. For the matrix and conventional approaches, the corresponding inverse problems have the form:

$$\mathsf{K}\,\mathbf{V}'_0 = \mathbf{R}' \tag{35}$$

$$\mathbf{K}\,\mathbf{V}'_0 = R' \tag{36}$$

where the matrix $\mathsf{K}$ is given by Eq. (30) and the row vector $\mathbf{K}$ is given by Eq. (34).

From the adjoint problem accomodating multiple observables Eqs. (28), (29), it is apparent that its *matrix* solution $\mathsf{V}^*$, and the resulting matrix $\mathsf{K}$ of the inverse problem Eq. (35) is not dependent on the residual $\mathbf{V}'$.

This is not the case for the *vector* solution $\mathbf{V}^*$ of the conventional adjoint problem Eqs. (32), (33) and resulting row vector $\mathbf{K}$ of the inverse problem Eq. (36). *The order of magnitude of* $\mathbf{K}$ *is proportional to the order of magnitude of the residual* $\mathbf{V}'$.

Thus, it can be expected that the availability of sensitivities of individual observables will enable a substantially faster convergence of the data assimilation process.

## 3. Conclusion

1. The matrix approach to the adjoint sensitivity analysis represents a natural way of treatment of individual observables in the process of variational of data assimilation.

2. The matrix approach gives an ability to select individual observables which are most sensitive to the model parameters to be retrieved; this is important both for planning of measurements and for variational assimilation of obtained data.

3. It can be expected that the application of the matrix approach to the variational assimilation of data may result in a substantial enhancement of convergence process and corresponding reduction of the number of iterations.

## Acknowledgements

The research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## 1. Matrix approach and vector of observables

First, we present the matrix approach to the adjoint sensitivity analysis to multiple observables, which was developed in (Ustinov, 2001) [1]. Let evolution of the system observed is described by the nonlinear initial value problem

$$\frac{d\mathbf{V}}{dt} + \mathcal{N}[\mathbf{V}, \mathbf{a}] = \mathbf{0} \tag{1}$$

$$\mathbf{V}|_{t_0} = \mathbf{V}_0 \tag{2}$$

Here $\mathcal{N}$ is a nonlinear operator that acts on the state vector $\mathbf{V}(t)$ in a way that, in general, depends on the vector of atmospheric parameters, $\mathbf{a}(t)$. Vector $\mathbf{V}_0$ consists of model parameters describing the initial state.

If the vector of parameters $\mathbf{a}$ experiences a perturbation $\mathbf{a}'$ then the forward problem can be linearized around its unperturbed, basic solution $\mathbf{V}(t)$ in the form

$$\frac{d\mathbf{V}'}{dt} + \mathsf{C}(t)\mathbf{V}'(t) = \mathbf{S}_e(t) \tag{3}$$

$$\mathbf{V}'|_{t_0} = \mathbf{S}_c \tag{4}$$

where the perturbations of all model parameters are contained in the matrix $\mathsf{C}(t)$ and in the right-hand term vectors $\mathbf{S}_e(t)$ and $\mathbf{S}_c$.

Let the $m-$vector of observables $\mathbf{R}$ be a result of action of some nonlinear operator $\mathcal{W}_e$ on the state vector $\mathbf{V}(t)$ over the interval $[t_0,\ t_1]$ where Eq.3 is integrated:

$$\mathbf{R}_e = \mathcal{W}_e[\mathbf{V}(t), t] \tag{5}$$

Similarly, the observables that are obtained at the end of the interval $[t_0,\ t_1]$ are considered as a result of action of some nonlinear operator $\mathcal{W}$ on the final value of the state vector $\mathbf{V}(t_1)$:

$$\mathbf{R}_c = \mathcal{W}_c[\mathbf{V}(t_1)] \tag{6}$$

## References


[1] Ustinov, E.A., 2001. Adjoint sensitivity analysis of atmospheric dynamics: Applications to the case of multiple observables. *J. Atmos. Sci.*, **58**, 3340–3348.

[2] Talagrand, O., and P. Courtier, 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I. Theory. *Ouart. J. Roy. Met. Soc.*, **113**, 1311–1328.

[3] O.Talagrand, 1997. Assimilation of observations, an introduction. *J. Meteorol. Soc. Japan* **75**, 191–209.