Practical global oceanic state estimation

Uriel Zajaczkovski

SIO219: SOSE seminar

Winter 2014

based on:

- Practical global oceanic estate estimation, C. Wunsch & P. Heimbach (2007)
- The Ocean Circulation Inverse Problem, C. Wunsch (1996)

Fundamental assumptions

(1) There is no proof of the numerical convergence of any known GCM in the limit of arbitrarily fine resolution.

(2) Oceanic observation errors, control elements, and model state variables are sufficiently close to Gaussian to justify the use of a minimum variance estimator.

(3) It is possible to use oceanographic data of arbitrary type.

The algebra of the problem

$$\mathbf{E(t)x(t) + n(t) = y(t)} \qquad [1]$$

E(t) design matrix of arbitrary size M x N

x(t) is the model state vector at discrete time t

n(t) is a stochastic noise vector such that:

$$< \mathbf{n}(t) >= 0, \quad < \mathbf{n}(t)\mathbf{n}(t')^T >= \delta_{tt'}\mathbf{R}(t) \qquad [2]$$

The algebra of the problem

$$\mathbf{x}(\mathbf{t}) = \mathbf{L}(\mathbf{x}(\mathbf{t}-\boldsymbol{\Delta}\mathbf{t}), \mathbf{B}(\mathbf{t}-\boldsymbol{\Delta}\mathbf{t})\mathbf{q}(\mathbf{t}-\boldsymbol{\Delta}\mathbf{t}), \boldsymbol{\Gamma}(\mathbf{t}-\boldsymbol{\Delta}\mathbf{t}) \times \mathbf{u}(\mathbf{t}-\boldsymbol{\Delta}\mathbf{t}), \mathbf{t}-\boldsymbol{\Delta}\mathbf{t}) \qquad [3]$$

L is a non linear operator (for our purposes is computer code).

q(t) are known sources, sinks and boundary conditions.

u(t) is the (unknown) control vector with moments:

$$< \mathbf{u}(\mathbf{t}) >= \mathbf{0}, \quad < \mathbf{u}(\mathbf{t})\mathbf{u}(\mathbf{t}')^{\mathbf{T}} >= \delta_{\mathbf{tt}'}\mathbf{Q}(\mathbf{t}) \qquad [4]$$

The cost function

$$J = \sum_{t=1}^{t_f} [\mathbf{y(t)} - \mathbf{E(t)x(t)}]^{\mathbf{T}} \mathbf{R(t)}^{-\mathbf{1}} [\mathbf{y(t)} - \mathbf{E(t)x(t)}] \tag{5}$$

$$+ [\mathbf{x_0} - \mathbf{x(0)}]^{\mathbf{T}} \mathbf{P(0)}^{-\mathbf{1}} [\mathbf{x_0} - \mathbf{x(0)}]$$

$$+ \sum_{t=0}^{t_f - 1} \mathbf{u(t)}^{\mathbf{T}} \mathbf{Q(t)}^{-\mathbf{1}} \mathbf{u(t)}$$

$$\tilde{\mathbf{x}}(0) = \mathbf{x_0}, \quad < (\mathbf{x_0} - \mathbf{x(0)})\mathbf{x_0} - \mathbf{x(0)}))^{\mathbf{T}} >= \mathbf{P(0)} \tag{6}$$

$$< \mathbf{n(t)n(t')}^{\mathbf{T}} >= \delta_{\mathbf{tt'}} \mathbf{R(t)}$$

$$< \mathbf{u(t)u(t')}^{\mathbf{T}} >= \delta_{\mathbf{tt'}} \mathbf{Q(t)}$$

Unconstrained least-squares (or the "adjoint method")

$$J' = J - 2\sum_{t=1}^{t_f} \mu(\mathbf{t})^{\mathbf{T}}[\mathbf{x}(\mathbf{t}) - \mathbf{L}(\mathbf{x}(\mathbf{t-1}),$$ [7]

$$\mathbf{B}(\mathbf{t} - \mathbf{\Delta t}) \times \mathbf{q}(\mathbf{t-1}), \mathbf{\Gamma}(\mathbf{t-1})\mathbf{u}(\mathbf{t-1}), \mathbf{t-1})]$$

Minimization of J leads to a set of nonlinear simultaneous equations (the normal equations):

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{u}(t)} = \mathbf{Q}(t)^{-1}\mathbf{u}(t) + \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\ \mathbf{Bq}(t),\ \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{u}(t)}\right)^{\mathrm{T}} \times \boldsymbol{\mu}(t+1) = \mathbf{0}, \quad 0 \leq t \leq t_f - 1$$ [8]

$$\frac{1}{2}\frac{\partial J'}{\partial \boldsymbol{\mu}(t)} = \mathbf{x}(t) - \mathbf{L}[\mathbf{x}(t-1),\ \mathbf{Bq}(t-1),\ \boldsymbol{\Gamma}\mathbf{u}(t-1)] = \mathbf{0}, \quad 1 \leq t \leq t_f$$ [9]

Unconstrained least-squares

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(0)} = \mathbf{P}(0)^{-1}[\mathbf{x}(0) - \mathbf{x}_0] + \left(\frac{\partial \mathbf{L}(\mathbf{x}(0),\ \mathbf{B}\mathbf{q}(0),\ \boldsymbol{\Gamma}\mathbf{u}(0))}{\partial \mathbf{x}(0)}\right)^{\mathrm{T}} \boldsymbol{\mu}(1) = \mathbf{0} \tag{10}$$

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(t)} = \mathbf{E}(t)^{\mathrm{T}}\mathbf{R}(t)^{-1}[\mathbf{E}(t)\mathbf{x}(t) - \mathbf{y}(t)] - \boldsymbol{\mu}(t) + \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\ \mathbf{B}\mathbf{q}(t),\ \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{x}(t)}\right)^{\mathrm{T}} \times \boldsymbol{\mu}(t+1) = \mathbf{0}, \quad 1 \le t \le t_f - 1 \tag{11}$$

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(t_f)} = \mathbf{E}(t_f)^{\mathrm{T}}\mathbf{R}(t_f)^{-1}[\mathbf{E}(t_f)\mathbf{x}(t_f) - \mathbf{y}(t_f)] - \boldsymbol{\mu}(t_f) = \mathbf{0} \tag{12}$$

Practical issues

ECCO configuration:

1 degree spatial resolution and 23 vertical levels

Number of elements of the state vector x(t)?

5.3 million elements

Total number of elements after time stepping every hour for 13 years (1992-2004)?

2.1×10^9 million elements!

∴ The number of equations is too large to be solved directly

Solve by iteration

Integrate Eq. (9) forward in time to produce a first estimate of **x**(t) using **u**(0) = 0 as a first guess:

$$\frac{1}{2}\frac{\partial J'}{\partial \boldsymbol{\mu}(t)} = \mathbf{x}(t) - \mathbf{L}[\mathbf{x}(t-1),\ \mathbf{Bq}(t-1),\ \boldsymbol{\Gamma}\mathbf{u}(t-1)] = \mathbf{0}, \quad 1 \le t \le t_f \tag{9}$$

Then integrate Eq. (11) backwards in time to produce a first estimate of μ(t):

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(t)} = \mathbf{E}(t)^{\mathrm{T}}\mathbf{R}(t)^{-1}[\mathbf{E}(t)\mathbf{x}(t) - \mathbf{y}(t)] - \boldsymbol{\mu}(t) + \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\ \mathbf{Bq}(t),\ \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{x}(t)}\right)^{\mathrm{T}} \times \boldsymbol{\mu}(t+1) = \mathbf{0}, \quad 1 \le t \le t_f - 1 \tag{11}$$

Solve by iteration

This is feasible as long as we know the partial derivatives:

$$\left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\ \mathbf{Bq}(t),\ \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{u}(t)}\right)^{\mathrm{T}}$$

$$\left(\frac{\partial \mathbf{L}(\mathbf{x}(t), \mathbf{Bq}(t), \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{x}(t)}\right)^{\mathrm{T}}$$

But these are the partial derivatives of J with respect to the problem parameters, so a quasi-Newton method can be used to reduce the value of J.

Recipe:

* Integrate forward (with an initial guess) and backward in time
* Reduce the value of J using a quasi-Newton method.
* Modify the problem parameters
* Integrate forward and backward again (an iteration)

Summary

The global oceanographic problem is at the present time focused primarily on smoothing rather than forecasting.

The method of Lagrange multipliers is used to pose the problem as one of unconstrained least-squares minimization.

An automatic differentiation tool is used to calculate the so-called adjoint code of the GCM.

Major problems today lie less with the numerical algorithms (least-squares problems can be solved by many means) than with the issues of data and model error.

The relationship of the weight matrices to the true error covariances remains uncertain, the solutions should be regarded as least-squares solutions, rather than minimum variance ones.