Correlation Modeling using a Generalized diffusion Equation

- Weaver & Courtier, 2001: QJRMS, 127, 1815 - 1846

$\underline{\underline{B}}$: background error covariance matrix
Symmetric $(L \times L)$ $L \rightarrow$ dimension of x

- Split $\underline{\underline{B}}$ into two parts:
  (i) block-diagonal univariate auto-covariances
  (ii) multivariate cross-covariances

- Partition B into balanced component and unbalanced component

(Derber & Bouttier, 1999: Tellus 51A, 195-221)

$$\underline{\underline{B}} = \underline{\underline{K}}_b \, \underline{\underline{B}}_u \, \underline{\underline{K}}_b^T$$

$\underline{\underline{B}}_u$: unbalanced univariance covariance matrix

- For now on, we will concern ourselves only with $\underline{\underline{B}}_u$

- $\underline{\underline{B}}_u$ can be further factorized as

$$\underline{\underline{B}}_u = \underline{\underline{\Sigma}} \, \underline{\underline{C}} \, \underline{\underline{\Sigma}} \qquad (31)$$

$\underline{\underline{\Sigma}}$: diagonal matrix of background standard deviations

$\underline{\underline{C}}$: symmetric matrix of background error correlation

- Recall $\underline{\underline{B}} = \underline{\underline{B}}^{1/2} \underline{\underline{B}}^{T/2}$

where $\underline{\underline{B}}^{1/2} = \underline{\underline{K_b}} \underline{\underline{\Sigma}} \underline{\underline{C}}^{1/2}$ (32)

$\underline{\underline{C}} = \underline{\underline{C}}^{1/2} \underline{\underline{C}}^{T/2}$

and $\delta \underline{x} = \underline{\underline{K_b}} \underline{\underline{\Sigma}} \underline{\underline{C}}^{1/2} \underline{v}$

An aside ...

- Consider a simple 1-D example assuming Gaussian statistics
  - normal Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

μ = mean of x
σ = standard deviation

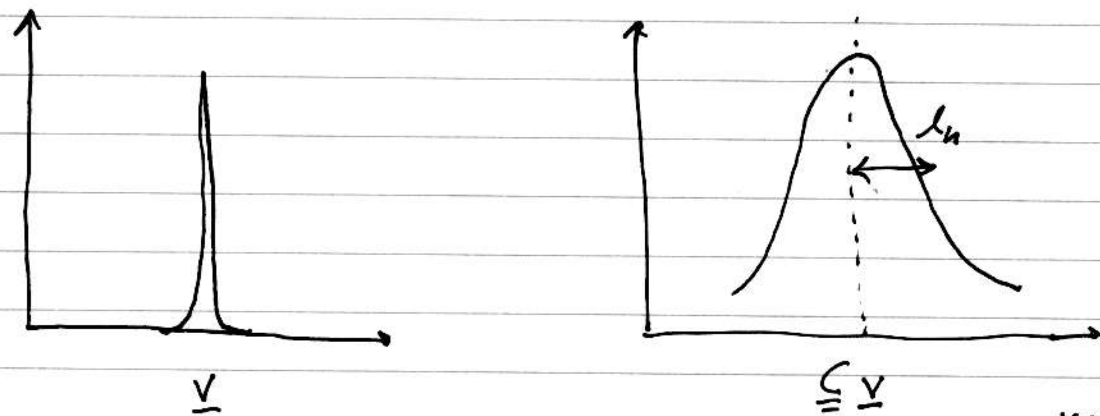
variance: $$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \underbrace{\int_{-\infty}^{\infty} (x-\mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}$$

This is the solution of a differential equation

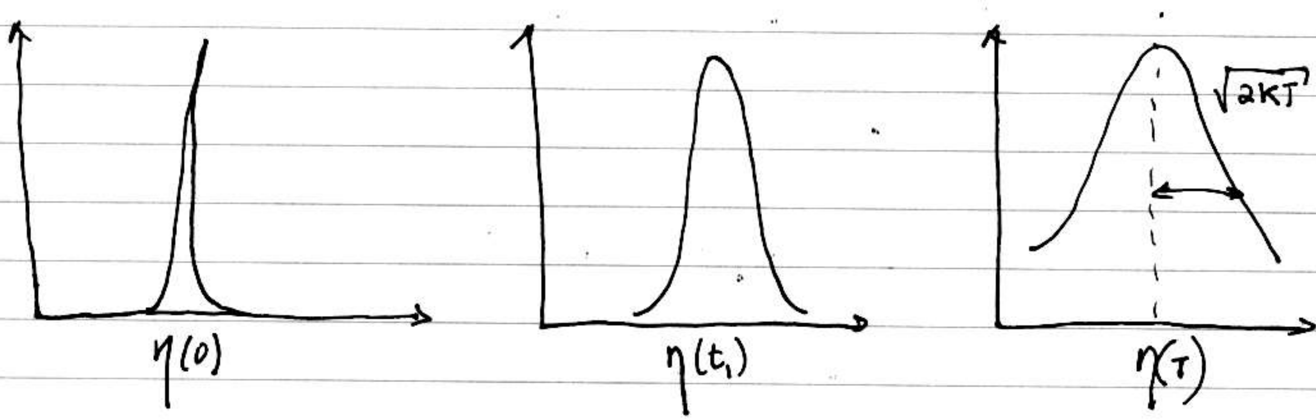
$$\left.\frac{\partial \eta}{\partial t} - \frac{\sigma}{2T} \frac{\partial^2 \eta}{\partial x^2} = 0 \right\} \text{ diff eq}$$

over the interval $t = [0, \tau]$ with $\eta(0) = (x-\mu)^2$

<u>Result:</u> A covariance function (or matrix) can be represented equivalently as the solution of an associated diffusion eq.



K: diffusion coef



- Let's assume that $\underline{\underline{C}}$ is a isotropic Gaussian spatial correlation function (matrix):

$$\underline{\underline{C}} \equiv C(\underline{r}, \underline{r}') = \exp\left(-|r-r'|^2/L^2\right)$$

position vector

$$\underline{u} = [u_i] = \underline{\underline{C}} v$$

$$u_i \equiv \int e^{-|\underline{r}-r'|^2/L^2} v(\underline{r}')\, d\underline{r}' \equiv u(\underline{r})$$

→ u_i can be found by integrating $\underline{v}$ with a diffusion equation.

- For a 2D field

$$\frac{\partial \eta}{\partial t} = k \nabla^2 \eta \qquad (33)$$

$$\eta(T) = \underline{\underline{L}}\, \eta(0)$$

$$\underline{\underline{L}} = \left\{ \underline{\underline{I}} + K \Delta t\, D \right\}^M$$

where $T = M \Delta t$.

D: discretized Laplacian operator

- Factorized L as $\underline{\underline{L}} = \underline{\underline{L}}^{1/2} \underline{\underline{L}}^{T/2}$

$$\underline{\underline{L}}^{1/2} = \left\{ \underline{\underline{I}} + K \Delta t\, D \right\}^{M/2} \qquad (34)$$

represent the solution of (33) at $t = T/2 = M\frac{\Delta t}{2}$ (half the # of diffusion equation timesteps)

- So $\underline{\underline{C}} = \underline{\underline{\Lambda}}\,\underline{\underline{L}}\,\underline{\underline{\Lambda}}$

$$\underline{\underline{C}}^{1/2} = \underline{\underline{\Lambda}}\,\underline{\underline{L}}^{1/2} \qquad (35)$$

where Λ is a matrix of normalization coefficients required to convert $\underline{\underline{L}}$ into a correlation matrix (i.e. with a range ± 1).

Where are we so far?

- We want to minimize $J(x)$ in (30) in v-space
- Having found $\underline{v}(0)$ we need to transform back to x-space:

$$\begin{aligned} \delta \underline{x}(0) &= \underline{\underline{B}}^{1/2}\,\underline{v}(0) \\ &= \underline{\underline{K}}_b\,\underline{\underline{\Sigma}}\,\Lambda\,\underline{\underline{L}}^{1/2}\,\underline{v}(0) \end{aligned} \qquad (36)$$

where $\underline{\underline{L}}^{1/2}$ represents the solution of the pseudo-diffusion (33) for $t = [0, T/2]$ with $\eta(0) = \underline{v}(0)$

- Vertical correlations can be handed in the same way by solving a vertical diffusion equation by assuming that $\underline{\underline{C}}$ is separable in the horizontal and vertical directions

$$\underline{\underline{C}}^{1/2} = \underline{\underline{\Lambda}}\,\underline{\underline{L}}_v^{1/2}\,\underline{\underline{L}}_h^{1/2}$$

where $\underline{\underline{L}}_v$ is a matrix operator representing solution of a discretized vertical diffusion equation

$$\frac{\partial \phi}{\partial t} = K_v \frac{\partial^2 \phi}{\partial z^2} \qquad (37)$$

- So $\underline{\underline{C}}^{1/2} \underline{v}(0) = \underline{\underline{\Lambda}}\, L_v^{1/2} \underline{\underline{L}}^{1/2} \underline{v}(0)$

(i) choose $\eta(0) = \underline{v}(0)$
(ii) Integrate (33) for $t = [0, T/2] \rightarrow \eta(T/2)$

(iii) Choose $\phi(0) = \eta(T/2)$
(iv) integrate (37) for $t = [0, T_v/2]$
(v) Multiply $\phi(T_v/2)$ by Λ

- The horizontal (l_h) and (l_v) correlation length are determined by

$$l_h^2 = 2 K_h T_h$$
$$l_v^2 = 2 K_v T_v$$

Determination of the Normalization factors Λ

- For isotropic correlations $\underline{\underline{C}}$ considered here, $\underline{\underline{\Lambda}}$ is simply a diagonal matrix and $\underline{\underline{\Lambda}}\, \underline{\underline{L}}_v^{1/2} \underline{\underline{\Lambda}}$ would simply be a matrix with 1's along diagonal (i.e. correlations of each field with itself at each grid point).

(i) An exact method
- To find Λ we therefore evaluate

$$\underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \underline{e} = \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_{l-1} \\ e_l \\ e_{l+1} \\ \vdots \\ e_L \end{matrix}$$

$\rightarrow$ l^{th} column of $\underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2}$ for $l = 1, \ldots, L$

The inverse square root of the corresponding l^{th} element of $\underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \underline{e}$ would represent the l^{th} diagonal element of Λ

$\longrightarrow$ computationally prohibitive!

(ii) A randomization method

- A cheaper alternative to the exact method
- Let $\underline{u}$: a Gaussian random vector with zero mean and unit variance

$$\langle \underline{u} \rangle = 0 \; ; \qquad \langle \underline{u}\,\underline{u}^T \rangle = \underline{\underline{I}}$$

- Now consider

$$\underline{\tilde{u}} = \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \underline{u}$$

$$\Rightarrow \quad \langle \underline{\tilde{u}} \rangle = \langle \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} u \rangle = \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \langle \underline{u} \rangle = 0$$

$$\langle \underline{\tilde{u}}\,\underline{\tilde{u}}^T \rangle = \langle \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \underline{u}\,\underline{u}^T \underline{\underline{L}}_h^{T/2} \underline{\underline{L}}_v^{T/2} \rangle$$

$$= \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \langle \underline{u}\,\underline{u}^T \rangle \underline{\underline{L}}_h^{T/2} \underline{\underline{L}}_v^{T/2}$$

$$= \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \underline{\underline{L}}_h^{T/2} \underline{\underline{L}}_v^{T/2}$$

Therefore the outer-product $\tilde{\underline{u}}\tilde{\underline{u}}^T$ of an ensemble of vectors $\tilde{\underline{u}}$ will yield an estimate of $(\underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2})(L_v^{1/2} L_h^{1/2})^T$ which is the square of the desired operator.

- Practically: choose an ensemble of Q random vectors $\underline{u}_i$ and we compute $\tilde{u}_i = \underline{\underline{L}}_v^{1/2} \underline{\underline{L}}_h^{1/2} \cdot \underline{u}_i$

- Then compute the diagonal elements of the ensemble outer-product

$$\frac{1}{Q} \sum_{i}^{Q} \tilde{\underline{u}}_i \tilde{\underline{u}}_i^T = \gamma$$

- The inverse square root of the diagonal elements of γ are the required diagonal elements of Λ.

The Complete IS4DVAR Algorithm for ROMS

- Compute/choose background error standard deviations: Σ
- Choose l_h and l_v, and then compute the normalization coefficients for the background error correlation: Λ
- Choose $\underline{x}_0(0) = \underline{x}_b(0)$

Outer-Loop : $n=1,p$

(1) integrate NLROMS, $t=[0,\tau]$ and save $\underline{x}_n(t)$
(2) Choose: $\delta\underline{x}_0(0) = \underline{v}(0) = 0$

Inner-Loops: $m=1,q$

(a) Integrate TL ROMS, $t=[0,\tau]$ and evaluate $J(\underline{x})$ using (27)
(b) Integrate ADROMS, $t=[\tau,0]$ forced by $-\underline{\underline{H}}_i^T \underline{\underline{O}}^{-1}(\delta\underline{x}_i - \underline{d}_i)$ to yield $\left(\partial J_o/\partial \delta x(0)\right) = -\lambda_m(0)$

(c) Compute $\left(\partial J_o/\partial v(0)\right)_m = \underline{\underline{B}}^{T/2}(-\lambda_m(0))$

$$= \underbrace{\underline{\underline{L}}_h^{T/2}\,\underline{\underline{L}}_v^{T/2}}_{\text{adjoint diff operator}}\,\underline{\underline{\Lambda}}^T\,\underline{\underline{\Sigma}}^T\left(-\underline{\lambda}_m(0)\right)$$

(d) Compute the total cost function gradient in v-space

$$\left(\frac{\partial J}{\partial \underline{v}(0)}\right) = \underbrace{\underline{v}_{m-1}(0)}_{\frac{\partial J_b}{\partial \underline{v}(0)}} + \left(\frac{\partial J_o}{\partial v(0)}\right)_m$$

$$\left(\frac{\partial J_o}{\partial V(o)}\right) = \left(\frac{\partial J_o}{\partial \delta x}\right)\left(\frac{\partial \delta x}{\partial \underline{V}(o)}\right) = B^{T/2}\left(\frac{\partial J_o}{\partial \underline{V}(0)}\right)$$

$$\delta \underline{x} = \underline{\underline{B}}^{1/2}\, \underline{V}$$

$$\frac{\partial\, \delta x}{\partial V} = \underline{\underline{B}}^{T/2}$$

(e) compute new donw-gradient descent direction

$$\underline{d}_m = -\left(\frac{\partial J}{\partial V(0)}\right) + \gamma_{m-1}\, \underline{d}_{m-1}$$

where $\gamma = 0$ for $m=1$ or $\mathrm{mod}\,(m, M) = 0$, $M \sim 5$
otherwise use (21)

(f) Choose trial step size α. (use α_p from previous inner-loop) and construct TL initial conditions in v-space

$$\underline{V}_m(0) = \underline{V}_{m-1}(0) + \alpha\, \underline{d}_m$$

(g) . Compute $\delta \underline{x}_m(0) = \underline{\underline{B}}^{1/2}\, \underline{V}_m(0) = \underbrace{\Sigma \Lambda L_V^{1/2} L_h^{1/2}}_{\text{TL diffusion operator}}\, \underline{V}_m(0)$

(h) Integrate TL ROMS and compute the "optimum step size", α_p using the incremental form of (22)

(i) Compute the new $\underline{V}_m(0) = \underline{V}_{m-1}(0) + \alpha_p\, \underline{d}_m$ and $\delta \underline{x}_m(0) = \underline{\underline{B}}^{1/2}\, \underline{V}_m(0)$

Goto (a)

End of the inner loop

(3) Compute new NLROMS initial conditions

$$\underline{x}_n(0) = \underline{x}_{n-1}(0) + \delta\, \underline{x}_q(0)$$

GOTO (1)

END of the outer loop

Weak Constraint 4DVAR Using the Method of Representers

- French derivation Courtier (1997, QJRMS, 123, 2449-2461)
  Dual formulation of 4DVAR: Physical-space statistical ~~space~~ Analysis scheme (PSAS)

- In 3ADVAR & IS4DVAR: N-dimensional system
  M-observations - info M-dimensions
  $M \ll N$

  searching the full N-dimensions for solutions by constraining the null space (N-M dimensions that are not observed) using the background $\underline{x}_b$

- Much more efficient to search in M-space (obs space) than in N-space (full state space).
- This can be quite naturally and conveniently achieved using Green's functions and representer functions.
- We are going to follow Chua & Bennett (2001, Ocean Modelling, 3, 137-165)

Notation:

NL ROM with errors: $$\frac{d\underline{x}}{dt} = N(\underline{x}) + \underbrace{F(t)}_{\text{External forcing}} + \overbrace{f(t)}^{\text{forcing errors and errors in the dynamics}} \qquad (37)$$

subject to: $$\underline{x}(0) = \underbrace{\underline{I}}_{\text{initial conditions}} + \underbrace{\underline{i}}_{\text{initial condition errors}} \qquad (38)$$

$$\underline{x}_\Omega(t) = \underbrace{\underline{B}(t)}_{\text{boundary conditions}} + \underbrace{b(t)}_{\text{boundary condition error}} \qquad (39)$$

Observations: (consider obs at a single time t)

$$\underline{d}(t) = \underline{\underline{H}}\,\underline{x}(t) + \underbrace{\varepsilon}_{\text{obs error}} \qquad (40)$$

Time interval: $t = [0, \tau]$

Assume unbiased, random, uncorrelated errors

Hypotheses:

$$\langle \underline{f} \rangle = \langle \underline{i} \rangle = \langle \underline{b} \rangle = 0 \qquad (41)$$

M×N $\quad \langle \underline{f}(t)\,\underline{f}^T(t') \rangle = \underline{\underline{C}}_f(t,t')$ $\quad$ forcing and dynamics error covariance matrix

M×N $\quad \langle \underline{i}\,\underline{i}^T \rangle = \underline{\underline{C}}_i$ $\quad$ initial conditions error covariance matrix

$\langle \underline{b}(t)\,\underline{b}^T(t') \rangle = \underline{\underline{C}}_b(t,t')$ $\quad$ boundary conditions error covariance matrix

M×M $\quad \langle \underline{\varepsilon}\,\varepsilon^T \rangle = \underline{\underline{O}}$ $\quad$ observation error covariance matrix

$$\left.\begin{array}{l} \langle \underline{f}\,\underline{b}^T \rangle = \langle \underline{f}\,i^T \rangle = \langle \underline{i}\,b^T \rangle = 0 \\ \langle \underline{f}\,\varepsilon^T \rangle = \langle \underline{i}\,\underline{\varepsilon}^T \rangle = \langle \underline{b}\,\underline{\varepsilon}^T \rangle = 0 \end{array}\right\} \text{uncorrelated errors}$$

Cost function (Penalty function)

$$\mathcal{J}(\underline{x}) = \int_0^\tau\int_0^\tau \underline{f}^T(t)\, \underline{\underline{C}}_f^{-1}(t,t')\,\underline{f}(t')\,dt\,dt' + \underline{i}^T \underline{\underline{C}}_i^{-1} \underline{i}$$

$$+ \int_0^\tau\int_0^\tau \underline{b}^T(t)\, \underline{\underline{C}}_b^{-1}(t,t')\,\underline{b}(t')\,dt\,dt' + \underline{\epsilon}^T \underline{\underline{O}}^{-1} \underline{\epsilon} .$$

$$= \int_0^\tau \left(\frac{d\underline{x}}{dt} - N(\underline{x}) - \underline{F}(t)\right)^T \overbrace{\int_0^\tau \underline{\underline{C}}_f^{-1}(t,t')\left(\frac{d\underline{x}}{dt'} - N(\underline{x}) - \underline{F}(t')\right) dt'}^{\underline{\lambda}(t)}\, dt$$

$$+ \left(\underline{x}(0) - \underline{I}\right)^T \underline{\underline{C}}_i^{-1}\left(\underline{x}(0) - \underline{I}\right) + \int_0^\tau\int_0^\tau \left(\underline{x}_\Omega(t) - \underline{B}(t)\right)^T \underline{\underline{C}}_b^{-1}\left(\underline{x}_\Omega(t') - \underline{B}(t')\right) dt'\,dt$$

$$+ \left(\underline{d} - \underline{\underline{H}}\,\underline{x}(t)\right)^T \underline{\underline{O}}^{-1}\left(\underline{d} - \underline{\underline{H}}\,\underline{x}(\tau)\right) \qquad (42)$$

Consider the first-variation $\delta\mathcal{J}(\underline{x})$ of $\mathcal{J}(\underline{x})$ resulting from a change $\delta\underline{x}(t)$:

$$\delta\mathcal{J}(\underline{x}) \simeq 2\int_0^\tau \left(\frac{\partial \delta\underline{x}}{dt} - \left(\frac{\partial N}{\partial \underline{x}}\right)\delta\underline{x}\right)^T \underline{\lambda}(t)\,dt + 2\,\delta\underline{x}^T(0)\, \underline{\underline{C}}_i^{-1}\left(\underline{x}(0) - \underline{I}\right)$$

$$+ 2\int_0^\tau\int_0^\tau \delta\underline{x}_\Omega^T(t)\, \underline{\underline{C}}_b^{-1}\left(\underline{x}_\Omega(t') - \underline{B}(t')\right) dt'\,dt$$

TLROMS

$$- 2\,\delta\underline{x}^T(\tau)\,\underline{\underline{H}}^T \underline{\underline{O}}^{-1}\left(\underline{d} - \underline{\underline{H}}\,x(\tau)\right)$$

$$2\int_0^\tau \delta\underline{x}^T \overbrace{\left(-\frac{\partial\underline{\lambda}}{dt} - \left(\frac{\partial N}{\partial x}\right)^T \underline{\lambda}\right)}^{\text{AD ROMS}} dt +$$

$$2\int_0^\tau \frac{d}{dt}\left(\underline{\lambda}^T(t)\,\delta\underline{x}(t)\right) dt$$

$$\searrow\ 2\left[\delta\underline{x}^T(\tau)\,\underline{\lambda}(\tau) - \delta\underline{x}^T(0)\,\lambda(0)\right]$$

At the Extrema

$$\frac{\partial J}{\partial \underline{x}(t)} = 0 \quad \Rightarrow \quad -\frac{d\lambda}{dt} - \left(\frac{\partial N}{\partial \underline{x}}\right)^T \underline{\lambda} - \delta(t-\tau)\underline{\underline{H}}^T \underline{\underline{O}}^{-1}(d - \underline{\underline{H}}\,x(\tau))$$

$$\frac{\partial J}{\partial x(\tau)} = 0 \quad \Rightarrow \quad \lambda(\tau) = 0$$

$$\frac{\partial J}{\partial \underline{x}(0)} = 0 \quad \Rightarrow \quad \underline{\underline{C}}_i^{-1}(\underline{x}(0) - \underline{I}) - \underline{\lambda}(0) = 0$$

$$\underline{x}(0) = \underline{I} + \underline{\underline{C}}_i\,\underline{\lambda}(0)$$

AD B.C.

$$\frac{\partial J}{\partial x_\Omega(t)} = 0 \quad \Rightarrow \quad \int_0^\tau \underline{\underline{C}}_b^{-1}(t,t')\left(\underline{x}_\Omega(t') - B(t')\right] dt - \left[\left(\frac{\partial N}{\partial x}\,\underline{\lambda}^T\right] = 0$$

$$\therefore \quad \underbrace{\underline{x}_\Omega(t) = \underline{B}(t) + \int_0^\tau \underline{\underline{C}}_b \left[\left(\frac{\partial N}{\partial x}\right)^T \underline{\lambda}(t')\right]_\Omega dt'}_{\int_0^\tau \underline{\underline{C}}_b(t,t')\,\underline{\underline{C}}_b^{-1}(t,t'')\,dt = \delta(t'-t'')}$$

Recall $$\underline{\lambda}(t) = \int_0^\tau \underline{\underline{C}}_f^{-1}(t,t')\left(\frac{d\underline{x}}{dt} - N(\underline{x}) - F(t')\right) dt'$$

$$\Rightarrow \quad \frac{d\underline{x}}{dt} - N(\underline{x}) = \underline{F}(t) + \int_0 \underline{\underline{C}}_f(t,t')\,\underline{\lambda}(t')\,dt'$$

The nonlinear Euler-Lagrange Equations:

ADROMS
$$-\frac{d\underline{\lambda}}{dt} - \left(\frac{\partial N^P}{\partial \underline{x}}\right)^T \underline{\lambda} = \delta(t-\tau)\, \underline{\underline{H}}^T \underline{\underline{O}}^{-1} \left(\underline{d} - \underline{\underline{H}}\, \hat{\underline{x}}(\tau)\right) \qquad (43)$$

subject to: $\underline{\lambda}(\tau) = 0$ (44)

$\underline{\lambda}_\Omega(t) = \underline{B}^\dagger(t)$ (45)

adjoint boundary conditions as a by product of the Green's identity

$\int \phi A \psi \, dx = \int \psi A^\dagger \phi \, dx$

$\underline{x}^T \underline{\underline{A}} \underline{y} = \underline{y}^T \underline{\underline{A}}^T \underline{x}$

NLROMS
$$\boxed{\frac{d\hat{\underline{x}}}{dt} - N^P(\hat{\underline{x}}) = \underline{F}(t)} + \int_0^\tau \underline{\underline{C}}_f(t,t')\, \underline{\lambda}(t')\, dt' \qquad (46)$$

subject to $\hat{\underline{x}}(0) = \underline{I} + \underline{\underline{C}}_i\, \underline{\lambda}(0)$ (47)

$$\hat{\underline{x}}_\Omega(t) = B(t) + \int_0^\tau \underline{\underline{C}}_b \left[\left(\frac{\partial N^P}{\partial \underline{x}}\right)^T \underline{\lambda}(t')\right]_\Omega dt' \qquad (48)$$

- The solution of (43)-(48) yields the optimal (minimum variance) estimate $\hat{\underline{x}}$.
- Because these equations are nonlinear we need to solve them iteratively.

Green's Functions and Representer Functions

- Suppose for a minute that ROMS is linear → the Euler-Lagrange equations are linear
- Consider ADROMS (43) forced by a delta function

$$-\frac{d\underline{\alpha}_m}{dt} - \left(\frac{\partial N^P}{\partial x}\right)^T \underline{\alpha}_m = \delta_m \qquad (49)$$

a delta function at a single observation point in space and time ("point m")

$\underline{\alpha}_m$ is a Green's function

- Solutions of (43) can be seen as linear superpositions of Green's functions.

- Note that (46), now assumed linear, is forced by α_m (i.e. forcing linear ROMS (TL ROMS) by Green's function)

$$\frac{d\underline{r}_m}{dt} + \left(\frac{\partial N}{\partial \underline{x}}\right)\underline{r}_m = \underline{\underline{C}}_f \underline{\alpha}_m \qquad (50)$$

$$\underline{r}_m(0) = \underline{\underline{C}}_i \alpha_m(0)$$

$$\underline{r}_m(t) = \int_0^{\tau} \underline{\underline{C}}_b(t') \underline{\alpha}_m(t')\, dt'$$

$\underline{r}_m$ are representer functions

- The representers have the property that:

* $\{\underline{r}_m, x\} = x_m \longrightarrow$ The part of $\underline{x}$-space that is spanned by the observations at m

inner product

* $\underline{\underline{R}}\,\underline{\underline{R}}^T\underline{\underline{S}} \longrightarrow$ representer covariance

$$\delta\underline{x}(t) = \underline{\underline{R}}\,\delta\underline{x}(0)$$

$$\int \underline{\underline{R}}\,\underline{\underline{C}}_f\,\underline{\underline{R}}^T \delta\, dt$$

$$\langle \delta\underline{x}(t)\,\delta\underline{x}^T(t)\rangle = \langle \underline{\underline{R}}\,\delta\underline{x}(0)\,\delta\underline{x}^T(0)\,\underline{\underline{R}}^T\rangle$$
$$= \underline{\underline{R}}\,\underbrace{\langle \delta\underline{x}(0)\,\delta\underline{x}^T(0)\rangle}_{\underline{\underline{C}}}\,\underline{\underline{R}}^T$$
$$= \underline{\underline{R}}\,\underline{\underline{C}}\,\underline{\underline{R}}^T$$

The representers are covariance functions. How one variable relates to every other in the space spanned by the observation.

The $\underline{r}_m$ are orthogonal.

- The representer functions yields that part of x-space that it is spanned by the observations
- So, to restrict our search to observation space we can search for estimates that are linear combinations of the representers, r_m:

$$\underline{X}(t) = \underline{X}_F(t) + \sum_{m=1}^{M} \beta_m \underline{r}_m(t) \qquad (51)$$

$\underline{X}_F$: A "prior" or first-guess estimate
β_m : representer coefficients

- It can be shown (Bennett, 2002: Inverse Modeling of the ocean and atmosphere, p 20-21):

$$(\underline{\underline{V}} + \underline{\underline{O}})\underline{\beta} = \underline{d} - \underline{\underline{H}}\,\underline{x}_f(t) \qquad (52)$$

where $\underline{\underline{V}} = (\underline{r}_1, \underline{r}_2, \underline{r}_3, \ldots, r_m)$ is the representer matrix an β is a vector of representer coefficients.

Indirect Representer Method

(51) and (52) implie that we must solve for β and $\underline{V}$. that is, we need to run the TLM and ADM for each observations.
$\Rightarrow$ very expensive!

However, it can also be shown (Bennett, 2002, p 60)

$$\beta = \underline{\underline{O}}^{-1} (d - \underline{\underline{H}} \hat{\underline{x}})$$

so (43) becomes

$$-\frac{d\lambda}{dt} - \left(\frac{\partial N}{\partial \underline{x}}\right)^T \underline{\lambda} = \delta(t-\tau) \underline{\underline{H}}^T \beta$$

- Rewrite (52) as:

$$\underline{\underline{P}}\, \underline{\beta} = \underline{d} - \underline{\underline{H}}\, x_F(\tau) \qquad (53)$$

where $P = \underline{\underline{V}} + \underline{\underline{O}}$

(53) can be solved by minimizing

$$\underline{\underline{P}}\, \psi - (\underline{d} - \underline{\underline{H}}\, \underline{x}_F(\tau)) \qquad (54)$$

- For any arbitrary vector ψ

$$\underline{\underline{V}}\, \psi = \int_0^\tau \underline{\underline{R}}\, \underline{\underline{C}}_f R^T \dots dt$$

$$\underline{\underline{P}}\, \psi = \underline{\underline{V}}\, \psi + \underline{\underline{O}}\, \psi$$

calculate this by equivalently integrating ADROMS and TLROMS

Then use conjugate gradient.

The Prior Estimate or First-Guess, $\underline{x}_F(t)$

To maintain linearity and operator symmetry, $x_F(t)$ is usually computed from a linerized of ROMS

~~[illegible]~~:

NLROMS: $$\frac{d\underline{x}}{dt} = N(x) + F(t)$$

Let ~~[illegible]~~ ~~[illegible]~~ $x_F = \underline{x} + \delta\underline{x}$

$$\therefore \quad \underbrace{\frac{d\underline{x}_p}{dt} + \frac{d\delta x}{dt}}_{\frac{dx_F}{dt}} = N(x_p + \delta x) + F(t)$$

$$\frac{dx_F}{dt} = N(\underline{x}_p) + \left(\frac{\partial N}{\partial \underline{x}}\right)\delta\underline{x} + F(t) + \cancel{h.o.t}$$

$$\boxed{\frac{dx_F}{dt} \simeq N(\underline{x}_p) + \left(\frac{\partial N}{\partial \underline{x}}\right)(\underline{x}_F - \underline{x}_p) + F(t)}$$

The W4DVAR Algorithm (#ifdef W4DVAR)

- Choose $\underline{x}(0)$ (say $\underline{x}_b(0)$)
- Integrate NLROMS, $t=[0,\tau] \rightarrow \underline{x}_0(t)$

<u>Outer Loop</u>: $n=1,p$

(1) integrate RPROMS using $\underline{x}_{n-1}$, $t=[0,\tau] \rightarrow \underline{x}_{n_F}(t)$

(2) Choose firt-guess $\underline{\psi} = \underline{\underline{O}}^{-1}(\underline{d} - \underline{\underline{H}}\, \underline{x}_{n_F})$

<u>Inner loops</u>: $m=1,q$

(a) Integrate ADROMS, $t=[\tau,0]$ forced by $\underline{\underline{H}}^T \underline{\psi} \rightarrow \underline{\phi}_m(t)$

(b) Integrate TLROMS, $t=[0,\tau]$ forced by $\int_0^\tau \underline{\underline{C}}_F(t,t')\,\underline{\phi}(t')\,dt' \rightarrow \underline{\theta}_m(t)$

(c) Use conjugate gradient descent algorithm to compute new $\underline{\psi}$ that will minimize (54)

This is a different conjugate gradient algorithm to solve a system of linear equations.

GOTO (a)

<u>END of INNER-LOOP</u>

(3) Integrate ADROMS, $t=[\tau,0]$ forced by $\underline{\underline{H}}^T \underline{\beta} \rightarrow \underline{\lambda}_n(t)$

(4) Integrate RPROMS, $t=[0,\tau]$ forced by

$$F(t) + \int_0^\tau \underline{\underline{C}}_f(t,t')\,\underline{\lambda}_n(t')\,dt' \rightarrow \underline{x}_n$$

GOTO 1

<u>END OF OUTER LOOP</u>

The WADVAR Algorithm (revisited)

- Choose / compute error standard deviations for $\underline{i}$, $\underline{b}$, $\underline{f}$
- Choose correlation lengths l_h and l_v
- Compute normalization factors $\underline{\Lambda}$
- choose $\underline{X}(0)$ (say $\underline{X}_b(0)$)
- Integrate NLROMS, $t=[0,\tau] \rightarrow \underline{X}_0(t)$
- Integrate RPROMS, linearized about $\underline{X}_0(t) \rightarrow \underline{X}_{0_F}(t)$

Same initial conditions and forcing

Outer Loops : $n=1,p$

(1) Choose firstguess $\underline{\psi}_0 = \underline{\underline{O}}^{-1}(d - \underline{\underline{H}}\, \underline{X}_{n-1_F})$

Inner loops: $m=1,q$

(a) Integrate ADROMS, $t=[\tau,0]$ forced by $\underline{\underline{H}}^T \underline{\psi} \rightarrow \underline{\Phi}_m(t)$, (I.C. $\underline{\Phi}_m(\tau)=0$)

(b) Integrate TLROMS, $t=[0,\tau]$ forced by
$\int_0^\tau \underline{\underline{C}}_f(t,t')\,\phi_m(t')\,dt' \rightarrow \underline{\theta}_m(t)$; (I.C. $\underline{\theta}_m(0) = \underline{\underline{C}}_i\, \phi_m(0)$)

(c) Use conjugate gradient to compute new $\underline{\psi}_m$ that will minimize $(\underline{\theta}_m + \underline{\underline{O}}\,\underline{\psi}_{m-1} - (\underline{d} - \underline{\underline{H}}\,\underline{X}_{n-1_F}(\tau)))$ eq (54)

GO TO a

END OF INNER LOOP

(2) Integrate ADROM, $t=[\tau,0]$ forced by $\underline{\underline{H}}^T \beta_n \rightarrow \underline{\lambda}_n(t)$, (I.C. $\underline{\beta}_n = \underline{\psi}_q$)

(3) Integrate RPROM, $t=[0,\tau]$ forced by
$F(t) + \int_0^\tau \underline{\underline{C}}_f(t,t')\,\underline{\lambda}_m(t')\,dt' \rightarrow X_{n_F}(t)$; $(\underline{X}_{n_F}(0) = \underline{I} + \underline{\underline{C}}_i \underline{\lambda}_n(0))$

GO TO 1

END OF OUTER LOOP

Refiments

1) More efficient line search algorithm for IS4DVAR
   - Eliminate the additional CALL to TLROMS in the inner-loops

2) Better conjugate gradient algorithm for IS4DVAR
   - MIQN3 is no in parallel, obsolete nowadays
   - CONGRAD from CERFACS (Lanczos algorithm) → using singular vector directions

3) W4DVAR preconditioning of (54)?

4) Improve the linear stability of the inner loop by coarser resolutions, simplify physics, increased dissipation

5) Full suite of "Weaver-Courtier" correlation functions

6) Vertical projection by EOF's

7) Simultaneous corrections to T and S. (Ricci & Weaver, JPO)

8) Changes scales in outer-loops

(9) Swap RPROM for NLROMS?

(10) K_b?

(11) K_H, K_V, spatially dependent to give meaningful spatial variations in t_h and t_v

(12) Nonlocal observations

(13) Outlier processing and removal (B.C.)

(14) Data operators for different observations

(15) Optimal observation design

(16) Relaxation term to the RP ROMS

Current ROMS Assimilation Projects : 6 month target

(1) IAS → Brian, Hernan, Ralph, Manu, Andy, Julio
→ IS4DVAR/W4DVAR → SST, SSH, Cruise Data

(2) CODAE - California → Chris, Milena, Andy, Peter(?)
Dave Foley, Frank Schwing, Jim Doyle
Carl Wunsch, Patrick Heimbach
→ IS4DVAR/W4DVAR → SST, SSH, Cruise Data
AD_SENSITIVITY
Various metrics

(3) GLOBEC - CGOA → Manu, Andy, Al Hermann, Liz Dobbins,
Zack Powell

AD_SENSITIVITY

(4) EAC → Wilkin, Javier, Gordon, Hernan
IS4DVAR/W4DVAR SST, SSH, Cruise Data

(5) Lombok Strait → Hernan, Julia, Enrique, Weiqing,
(Manu, Andy)

(6) LaTTE → Wilkin, Gordon

(7) MURI → Wilkin
IS4DVAR/W4DVAR

(8) NASA, CalCOFI forecasting & hindcasting
Manu, Art
NLM predictability studies
Assimilate CALCOFI cruises

(9) CIMT, Monterey Bay → CENCOOS