Lecture 3:
Dual 4D-Var
Outline

• 4D-Var recap
• Dual 4D-Var
• The ROMS dual algorithms
• Weak constraint 4D-Var
Data Assimilation: Recap

Model solutions depends on $x_b(0)$, $f_b(t)$, $b_b(t)$, $h(t)$
Notation & Nomenclature: Recap

\[
x = \begin{bmatrix} T \\ S \\ \zeta \\ u \\ v \end{bmatrix} \quad z = \begin{bmatrix} x(0) \\ f(t) \\ b(t) \\ \eta(t) \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad d = \left( y - H(x^b) \right)
\]

State vector  Control vector  Observation vector

Prior

Innovation vector

Observation operator
Incremental Formulation: Recap

\( \delta z = (\delta x^T(0), \varepsilon_b^T(t), \varepsilon_f^T(t), \eta(t))^T \)

\( J = \frac{1}{2} \delta z^T D^{-1} \delta z + \frac{1}{2} \left( G \delta z - d \right)^T R^{-1} \left( G \delta z - d \right) \)

\( D = \text{diag}(B_x, B_b, B_f, Q) \)

\( f_b(t), B_f \)

\( b_b(t), B_b \)

Prior (background) error covariance

(Courtier et al., 1994)
Primal vs Dual Formulation: Recap

Vector of increments

Primal Space

Observation vector

Dual Space
The Solution: Recap

Analysis: \[ Z_a = Z_b + Kd \]

Gain (dual form):
\[ K = DG^T (GDG^T + R)^{-1} \]

Gain (primal form):
\[ K = (D^{-1} + G^T R^{-1} G)^{-1} G^T R^{-1} \]
Two Spaces: Recap

Gain (dual):

\[ K = D G^T (G D G^T + R)^{-1} \]
\[ = N_{\text{obs}} \times N_{\text{obs}} \]

Gain (primal):

\[ K = (D^{-1} + G^T R^{-1} G)^{-1} G^T R^{-1} \]
\[ = N_{\text{model}} \times N_{\text{model}} \]

\[ N_{\text{obs}} \ll N_{\text{model}} \]
Two Spaces: Recap

$G$ maps from model space to observation space

$G^T$ maps from observation space to model space
Primal Formulation: Recap

Analysis: \( z_a = z_b + Kd \)

Goal of 4D-Var is to identify:

\[
\delta z = Kd = (D^{-1} + G^T R^{-1} G)^{-1} G^T R^{-1} d
\]

Solve the equivalent linear system:

\[
(D^{-1} + G^T R^{-1} G)\delta z = G^T R^{-1} d
\]

by minimizing:

\[
J = \frac{1}{2} \delta z^T (D^{-1} + G^T R^{-1} G)\delta z - \delta z^T G^T R^{-1} d + \frac{1}{2} d^T R^{-1} d
\]

\[
= \frac{1}{2} \delta z^T D^{-1} \delta z + \frac{1}{2} (G \delta z - d)^T R^{-1} (G \delta z - d)
\]
Analysis: \[ z_a = z_b + Kd \]

Goal of 4D-Var is to identify:

\[ \delta z = Kd = DG^T (GDG^T + R)^{-1} d \]

Solve the equivalent linear system:

\[ (GDG^T + R)w = d; \quad \delta z = DG^T w \]

by minimizing:

\[ I(w) = \frac{1}{2} w^T (GDG^T + R)w - w^T d \]

then compute:

\[ \delta z = DG^T w \]

There is no physical significance attached to \( w \).
Conjugate Gradient (CG) Methods

Contours of $J$
Matrix-less Operations

There are no matrix multiplications!

\[ GDG^T \delta \]

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

\[ \text{GDG}^T \delta \]

Adjoint Model

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

$$GDG^T \delta$$

Adjoint Model

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{GDG}^T \delta$$

Adjoint Model

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

\[ (G^T \delta) \]

Covariance

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

$\text{GDG}^T \delta$

Covariance

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

$$GDG^T \delta$$

Tangent Linear Model

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

\[ \text{GDG}^T \delta \]

Tangent Linear Model

Zonal shear flow
Matrix-less Operations

There are no matrix multiplications!

$\text{GDG}^T \delta$

Zonal shear flow

A covariance
Matrix-less Operations

There are no matrix multiplications!

\[ \text{GDG}^T \delta \]

Tangent Linear Model

Zonal shear flow
**Dual 4D-PSAS Algorithm**

(defined \( W4DPSAS, w4dpsas\_ocean.h \))

\[ x_b(t), d \]

obs space

\[ G^T w \quad \text{ADROMS forced by } w \]

\[ D \]

\[ \text{Run ADROMS} \]

\[ D \]

\[ \text{Run TLROMS} \]

\[ \text{Conjugate Gradient Algorithm} \]

\[ \partial I / \partial w = (GDG^T + R)w - d \]

\[ G^T w_a \]

\[ \delta z_a \]

\[ \text{NLROMS, } z_a \]

\[ x_a(t) \]
Dual 4D-PSAS Algorithm
(define W4DPSAS, w4dpsas_ocean.h)

NLROMS, \( z_b \)

Choose \( w \)

Run ADROMS

D

Run TLROMS

Conjugate Gradient Algorithm

Run ADROMS

D

NLROMS, \( z_a \)
Weak Constraint 4D-Var

Nonlinear ROMS (NLROMS):

\[ \mathbf{x}(t_i) = M(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i)) \]

Nonlinear ROMS (NLROMS) with model error:

\[ \mathbf{x}(t_i) = M(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i), \epsilon(t_i)) \]

Model error prior: 0

Model error prior covariance: \( \mathbf{Q} \)

(No explicit time correlation in \( \mathbf{Q} \), but there is some in practice)

4D-Var control vector: \( \mathbf{z} = \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{f}(t) \\ \mathbf{b}(t) \\ \mathbf{\eta}(t) \end{bmatrix} \)

Correction for model error
Weak Constraint 4D-Var

Tangent linear ROMS (TLROMS):

\[ \delta x(t_i) = M(t_i, t_{i-1}) \delta u(t_{i-1}) \]

\[ \delta u(t_i) = \begin{bmatrix} \delta x(t_i) \\ \delta f(t_i) \\ \delta b(t_i) \\ \delta \eta(t_i) \end{bmatrix} \]

4D forcing for TLROMS

Strong constraint: \( \delta \eta(t_i) = 0 \)
Two Spaces

Gain (dual):

$$K = DG^T \left( GDG^T + R \right)^{-1}$$

Gain (primal):

$$K = \left( D^{-1} + G^T R^{-1} G \right)^{-1} G^T R^{-1}$$

$$N_{\text{obs}} \times N_{\text{obs}}$$

$$N_{\text{model}} \times N_{\text{model}}$$

$$N_{\text{obs}} \ll N_{\text{model}}$$
Two Spaces

Strong constraint:

\[ N_{\text{model}} = N_x + N_{\text{times}} \left( N_f + N_b \right) \]

Weak constraint:

\[ N_{\text{model}} = N_x + N_{\text{times}} \left( N_f + N_b + N_x \right) \]

Weak constraint is only practical in dual formulation of 4D-Var since \( N_{\text{obs}} \) is unaffected:

\[ K = DG^T \left( GDG^T + R \right)^{-1} \]

\[ \underbrace{N_{\text{obs}} \times N_{\text{obs}}} \]
Dual 4D-Var & Preconditioning

• While it appears that the dual formulation should be an easier problem to solve because of the considerably smaller dimension of the space involved, it has not been widely adopted because of practical barriers to convergence.
• Until recently preconditioning was a major hurdle to advances in the dual algorithm.
• We will review some recent progress in the framework of ROMS:
  (i) $R^{-1/2}$ preconditioning (PSAS – Courtier, 1997)
  (ii) MINRES vs CG (El Akkraoui and Gauthier, 2010)
  (iii) RPCG (Gratton and Tshimanga, 2009)
Dual 4D-Var: Naïve $R^{-1/2}$ Preconditioning

Analysis: $z_a = z_b + Kd$

Goal of 4D-Var is to identify:

$$\delta z = Kd = DG^T (GDG^T + R)^{-1} d$$

Solve the equivalent linear system:

$$(GDG^T + R)w = d; \quad \delta z = DG^T w$$

by minimizing:

$$I(w) = \frac{1}{2} w^T (GDG^T + R)w - w^T d$$

Preconditioning via the change of variable

$$v = R^{-1/2} w$$
Dual 4D-Var: CG with $R^{-1/2}$ Preconditioning

Lanczos formulation of conjugate gradient algorithm in observation space is used (congrad.F).

Dual formulation of gain matrix:

$$K = DG^T (GDG^T + R)^{-1}$$

Dual formulation of practical gain matrix:

$$\tilde{K}_k = DG^T R^{-1/2} \tilde{V}_k T_k^{-1} \tilde{V}_k^T R^{-1/2}$$

$\tilde{V}_k$ dual Lanczos vectors
Dual 4D-Var: Minimum Residual Method

Lanczos formulation of minimum residual algorithm in observation space is used (define MINRES).
Dual 4D-Var: Restricted Preconditioned CG (RPCG)

- Experience in primal 4D-Var has shown that preconditioning by $B^{-1/2}$ is very effective.
- Preconditioning by $B^{-1/2}$ can be enforced in dual 4D-Var by “restricting” the dual Lanczos vectors so that:

$$V_k = G^T \tilde{V}_k$$

- define RPCG
An Example: ROMS CCS

COAMPS forcing $f_b(t), B_f$

ECCO open boundary conditions $b_b(t), B_b$

Previous assimilation cycle $x_b(0), B_x$

30km, 10 km & 3 km grids, 30-42 levels

Veneziani et al (2009)
Broquet et al (2009)
Moore et al (2010)
Observations (y)

CalCOFI & GLOBEC

Ingleby and Huddleston (2007)

TOPP Elephant Seals

Data from Dan Costa

Argo
4D-Var Configuration

• Case studies for a representative case
  29 March - 4 April, 2003, 10km resolution.
• 1 outer-loop, 50 inner-loops
• 7 day assimilation window
• Prior $\mathbf{D}$: $\mathbf{x}$ $L_h=50$ km, $L_v=30$ m, $\sigma$ from clim
  \begin{align*}
  &\mathbf{f}\ L_\tau=300$km, $L_Q=100$km, $\sigma$ from COAMPS \\
  &\mathbf{b}\ L_h=100$ km, $L_v=30$ m, $\sigma$ from clim
\end{align*}
• Super observations formed
• Obs error $\mathbf{R}$ (diagonal):
  SSH 2 cm
  SST 0.4 C
  hydrographic 0.1 C, 0.01psu
Dual 4D-Var

(Gürol et al, 2013, QJRMS)
**Dual 4D-Var vs Primal 4D-Var**

- **Mid-Atlantic Bight, 7km, 2014 test case**
- 2 outer-loops
- 7 inner-loops
- 3 day windows
- SSH, SST, in situ, HF radar obs

- **CG iterations**
  - Dual 4D-Var is ~15-25% faster than primal 4D-Var in ROMS
  - Dual 4D-Var has more post-processing utility in ROMS
  - Only Dual 4D-Var supports the weak constraint option
Summary

• Strong and weak constraint 4D-Var, dual formulation:
  define W4DPSAS
  Drivers/w4dpsas_ocean.h

• Matrix-less iterations to identify cost function minimum using TLROMS and ADROMS

• RPCG: define RPCG
References

References

