Preconditioners for saddle point weak-constraint 4D-Var with correlated observation errors

Jemima M. Tabeart
University of Edinburgh
jemima.tabeart@ed.ac.uk
@jemimatabeart

With John Pearson
University of Edinburgh

Weak constraint 4D-Var

- Balancing information from observations and background over a time window
- Allow our model to be imperfect

$$\min_{\mathbf{x}\in\mathbb{R}^s} J(\mathbf{x}) = \left\|\mathbf{x}^{(0)} - \mathbf{x}_{\mathrm{b}}\right\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\sum_{j=0}^{n}\left\|\mathrm{H}_j\left(\mathbf{x}^{(j)}\right) - \mathbf{y}_j\right\|^2_{\mathbf{R}_j^{-1}} + \frac{1}{2}\sum_{j=1}^{n}\left\|\mathbf{x}^{(j)} - \mathbf{M}_j\left(\mathbf{x}^{(j-1)}\right)\right\|^2_{\mathbf{Q}_j^{-1}}$$

Saddle point formulation of weak-constraint data assimilation

Can re-write objective function in saddle point form (after linearising operators)

$$\begin{pmatrix} D & 0 & L \\ 0 & R & H \\ L^T & H^T & 0 \end{pmatrix} \begin{pmatrix} \delta\eta \\ \delta\mu \\ \delta x \end{pmatrix} = \begin{pmatrix} b \\ d \\ 0 \end{pmatrix}$$

$$D = \begin{pmatrix} B & & & & \\ & Q_1 & & & \\ & & Q_2 & & \\ & & & \ddots & \\ & & & & Q_n \end{pmatrix} \qquad R = \begin{pmatrix} R_0 & & & & \\ & R_1 & & & \\ & & R_2 & & \\ & & & \ddots & \\ & & & & R_n \end{pmatrix}$$

$$H = \begin{pmatrix} H_0 & & & & \\ & H_1 & & & \\ & & H_2 & & \\ & & & \ddots & \\ & & & & H_n \end{pmatrix} \qquad L = \begin{pmatrix} I & & & & & \\ -M_0 & I & & & & \\ & -M_1 & I & & & \\ & & -M_2 & I & & \\ & & & \ddots & \ddots & \\ & & & & -M_n & I \end{pmatrix}$$

Advantages and challenges of the saddle point problem for DA

Advantages:

- Parallelisable nature

Challenges:

- Typical assumption (1,1) block is hard, (2,1) block is easy (opposite in DA): Theory/heuristics don't always apply (invalidating standard assumptions)
- Designing preconditioners: prior work often assumes can invert (1,2) block exactly

Approximating L (the model)

$$L = \begin{bmatrix} I & & & & & \\ -M_0 & I & & & & \\ & -M_1 & I & & & \\ & & -M_2 & I & & \\ & & & & \ddots & \\ & & & & -M_n & I \end{bmatrix}$$

$$L^{-1} = \begin{bmatrix} I & & & \\ M_0 & I & & \\ M_0 M_1 & M_1 & \ddots & \\ M_0 M_1 M_2 & M_1 M_2 & & \ddots \\ \vdots & \vdots & & \\ M_0 \dots M_n & M_1 \dots M_n & M_n & I \end{bmatrix}$$

Standard approximations L

$$L_0 = \begin{bmatrix} I & & & & \\ & I & & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \end{bmatrix} \qquad L_1 = \begin{bmatrix} I & & & & \\ -I & I & & & \\ & -I & I & & \\ & & -I & \ddots & \\ & & & -I & I \end{bmatrix}$$

Proposed L

$$
L_M = \begin{bmatrix}
I &&&&&&& \\
-M_0 & I &&&&&& \\
& -M_1 & I &&&&& \\
&& \cancel{-M_2} & I &&&& \\
&&& \cancel{-M_3} & I &&& \\
&&&& -M_4 & I && \\
&&&&& \cancel{-M_5} & I & \\
&&&&&& -M_6 & I
\end{bmatrix}
$$

$k=3$

$k=4$

$$
L_M^{-1} = \begin{bmatrix}
I &&&&&&& \\
M_0 & I &&&&&& \\
\underline{M_0 M_1} & M_1 & I &&&&& \\
& 0 & 0 & I &&&& \\
&&& M_3 & I &&& \\
&&& \underline{M_3 M_4} & M_4 & I && \\
&&&& 0 & 0 & I & \\
&&&&&& M_6 & I
\end{bmatrix}
$$

Eigenvalues of $\mathbf{L}_M^{-1}\mathbf{L}\mathbf{L}^T\mathbf{L}_M^{-1}$

Assume: $\mathbf{M}_i \equiv \mathbf{M}$ for all subwindows.

Theorem

Let $\mathbf{M}$ be symmetric with eigenvalues in $\gamma \leq \lambda(\mathbf{M}) \leq \mu$. Then the eigenvalues of $\mathbf{L}\mathbf{L}^T$ are bounded above by

$$\lambda(\mathbf{L}\mathbf{L}^T) \leq 1 + \mu^2 + \max(2|\gamma|, 2|\mu|) \quad (1)$$

Theorem

Let $\mathbf{M}$ be symmetric with $\sigma(\mathbf{M}) \leq 1$ then for $k = 3$ we can bound the largest eigenvalue of $\mathbf{L}_M^{-1}\mathbf{L}\mathbf{L}^T\mathbf{L}_M^{-1}$ above by 8.

Numerical experiments

- $\mathbf{M}_i$ = Heat equation with zero Dirichlet boundary conditions
- $\mathbf{B}$ = spatial correlations ($s \times s$)
- $\mathbf{Q}_i$ = spatial correlations
- $\mathbf{R}_i$ = random entries with block structure ($p \times p$)
- $\mathbf{H}_i$ = random selection of half of variables
- $n + 1$ = number of observation times

Eigenvalues of $L_M^{-1} L L^T L_M^{-1}$

$n+1$	no units	No missing blocks	Upper bound	Largest eval
3	1500	0	-	1
4	1000	1	3	2.618
5	1500	1	4	3.7319
6	2000	1	5	4.791
7	1500	2	6	4.8928
8	2000	2	7	5.105
9	2500	2	7.5	5.5546
10	2000	3	8	5.5940
11	2500	3	8	5.6745
12	3000	3	8	5.8798
13	2500	4	8	5.8985

Table: Number of unit eigenvalues of $L_M^{-1} L L^T L_M^{-1}$ for $s = 500$ for k=3 (skipping every third block) for different number of observation subwindows. Each subwindow is the same length. Number of missing subwindows = floor(n/3). $i =$ remainder$(\frac{n+1}{\text{missing blocks}})$

Approximating the observation error covariance matrix

- $\mathbf{R}_{bl}$ - split into equally sized blocks
- $\mathbf{R}_{RR} = \mathbf{R} + \delta\mathbf{I}$ (increase all eigenvalues)
- $\mathbf{R}_{ME}$ increase all eigenvalues below a threshold T to be equal to T

Theorem

The eigenvalues of $\mathbf{R}_{RR}^{-1}\mathbf{R}$ *are given by*

$$\lambda(\mathbf{R}_{RR}^{-1}\mathbf{R}) = \frac{\lambda(\mathbf{R})}{\lambda(\mathbf{R}) + \delta}. \quad (2)$$

Theorem

The eigenvalues of $\mathbf{R}_{ME}^{-1}\mathbf{R}$ *are given by*

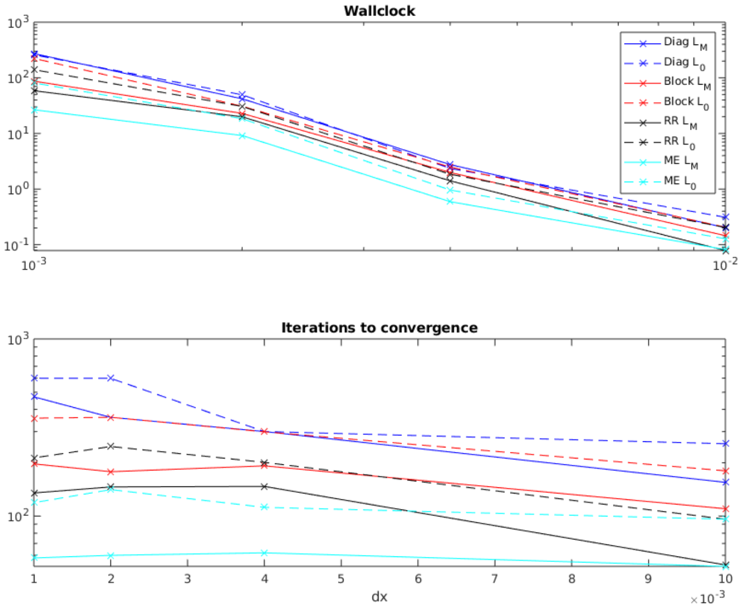
$$\lambda_i(\mathbf{R}_{ME}^{-1}\mathbf{R}) = \begin{cases} 1 & if\, \lambda_i > T \\ \frac{\lambda_i}{T} < 1 & if\, \lambda_i \leq T \end{cases} \quad (3)$$

Numerical experiments

$$P = \begin{pmatrix} D & 0 & 0 \\ 0 & \tilde{R} & 0 \\ 0 & 0 & \tilde{S} \end{pmatrix} \qquad \tilde{S} = L_M^T D^{-1} L_M$$

- Use exact $\mathbf{D}$
- Compare $\widetilde{\mathbf{L}} = \mathbf{L}_0$ and $\widetilde{\mathbf{L}} = \mathbf{L}_M$
- Compare $\widetilde{\mathbf{R}} = diag(\mathbf{R})$ with $\widetilde{\mathbf{R}} = \mathbf{R}_{bl}, \mathbf{R}_{RR}, \mathbf{R}_{ME}$
- Serial experiments using MINRES for increasing dimension n, fixed number of subwindows.

Wallclock time vs convergence



Conclusions

- Proposed a new preconditioner for $\mathbf{L}$ in the saddle point formulation of wc4D-Var
- Multiple approaches to precondition $\mathbf{R}$ when it is correlated
- Can bound the eigenvalues of the preconditioned $\mathbf{L}$ and $\mathbf{R}$ terms
- Iterations and wallclock time are improved compared to standard preconditioners

Conclusions

- Proposed a new preconditioner for $\mathbf{L}$ in the saddle point formulation of wc4D-Var
- Multiple approaches to precondition $\mathbf{R}$ when it is correlated
- Can bound the eigenvalues of the preconditioned $\mathbf{L}$ and $\mathbf{R}$ terms
- Iterations and wallclock time are improved compared to standard preconditioners

Ongoing work

- Bounds on preconditioned saddle point system
- Parallelisable $\widetilde{\mathbf{R}}$? Can we do better than naive diagonal approach?
- More realistic experiments: more complicated PDEs
  New preconditioners are parallelisable, but currently running in serial