CG vs MG solvers for diffusion-based correlation models in data assimilation

Philippe Toint (with Serge Gratton and Jean Tshimanga)

University of Namur, Belgium

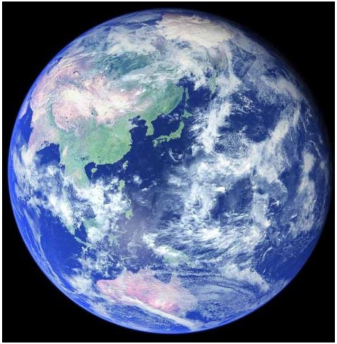
IMA, Birmingham, September 2012

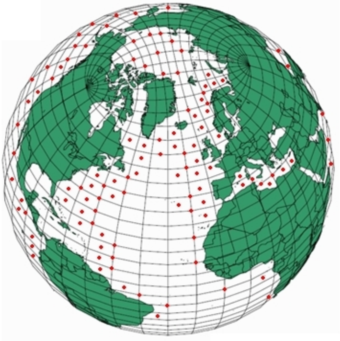
(Attempt to predict...

- tomorrow's weather
- the average ocean temperature next month
- the future gravity field
- the next ionospheric currents
- ...

Data assimilation in weather forecasting (2)

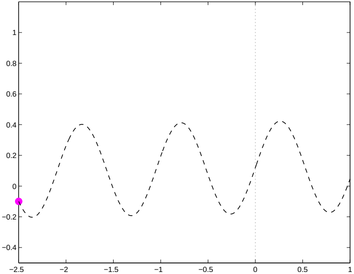
Data: température, wind, pression, . . . everywhere and at all times !





May involve more than 1.000.000.000 variables!

The principle:

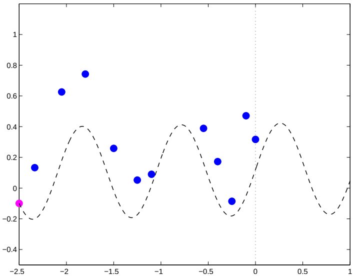


temperature/days

- Situation 2.5 days ago and "background" prediction

Data assimilation in weather forecasting (3)

The principle:

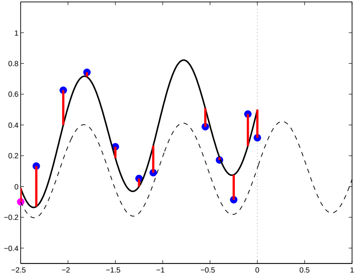


temperature/days

- Situation 2.5 days ago and “background” prediction
- Température for the last 2.5 days

The principle:

Minimize the error between model and past observations



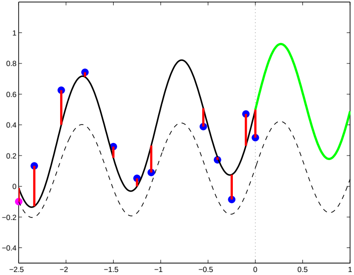
temperature/days

- Situation 2.5 days ago and "background" prediction
- Température for the last 2.5 days
- Run the model to minimize the gap between | model and observations

$$\min_{x_0} \frac{1}{2}\|x_0 - x_b\|^2_{B^{-1}} + \frac{1}{2}\sum_{i=0}^{N} \|\mathcal{HM}(t_i, x_0) - b_i\|^2_{R_i^{-1}}.$$

The principle:

Minimize the error between model and past observations



temperature/days

- Situation 2.5 days ago and "background" prediction
- Température for the last 2.5 days
- Run the model to minimize the gap between | model and observations
- Predict tomorrow's temperature

The 4D-VAR approach

$$\min_x \frac{1}{2}\|x - x_b\|^2_{B^{-1}} + \frac{1}{2}\sum_{i=0}^{N} \|\mathcal{H}\mathcal{M}(t_i, x) - b_i\|^2_{R_i^{-1}}.$$

- a weighted nonlinear least-squares problems
- $\Rightarrow$ a Gauss-Newton (linearization) approach
- $\Rightarrow$ iterately solve (level-1 iterations)

$$\min_x \frac{1}{2}\langle x - x_b, B^{-1}(x - x_b)\rangle + \frac{1}{2}\langle Hx - d, R^{-1}(Hx - d)\rangle$$

(a (very) large quadratic minimization problem)

Solving the 4D-VAR subproblem

Analytic solution:

$$(I + BH^T R^{-1} H)x = BH^T R^{-1} d$$

In pratice:

- use Conjugate Gradients (or other Krylov space solver)
- for a (very) limited number of level-2 iterations
- (with preconditioning, but not discussed here)

$\Rightarrow$ need products of the type

$$(I + BH^T R^{-1} H)v \quad \text{for a number of vectors } v$$

Focus here on how to compute Bv (B large)

A widely used approach (Derber + Rosati, 1989, Weaver + Courtier, 2001):

Spatial background correlation ≈ diffusion process

i.e.

Computing Bv
$\approx$
integrating a diffusion equation starting from the state v.

use p steps of an implicit integration scheme

(level-3 iteration, each involving a solve with B!!!)

The integration iteration

Define

$$\Theta_h = I + \frac{\mathcal{L}}{2p}\Delta_h$$

(Δ_h is the discrete Laplacian, $\mathcal{L}$ is the correlation length).
For each integration (z and p given)

1. $u_0 = \left(\mathrm{diag}(\Theta_h^{-p}\right)^{-1/2} z$ (diagonal scaling)
2. $u_\ell = \Theta_h^{-1} u_{\ell-1}$ $(\ell = 1, \ldots, p)$
3. $Bz = \left(\mathrm{diag}(\Theta_h^{-p}\right)^{-1/2} u_p$ (diagonal scaling)

Our question: how to solve $\Theta_h u_\ell = u_{\ell-1}$?

The integration iteration

Our question: how to solve $\Theta_h u_\ell = u_{\ell-1}$?

Carrier + Ngodock (2010):

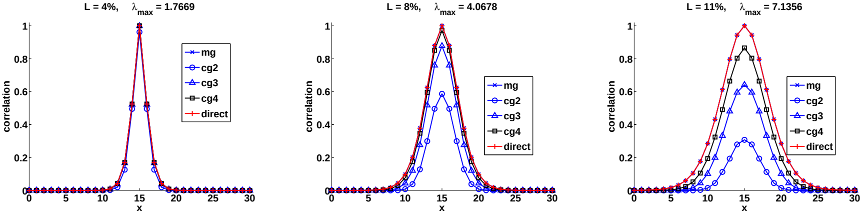
Implicit integration + CG is $\approx$ 5 times faster than explicit integration!

But:

- What about multigrid ??
- Is an approximate solution of the system (CG or MG) altering the spatial properties of the correlation?
- Inexact solves ?

Approximately diffusing a Dirac pulse

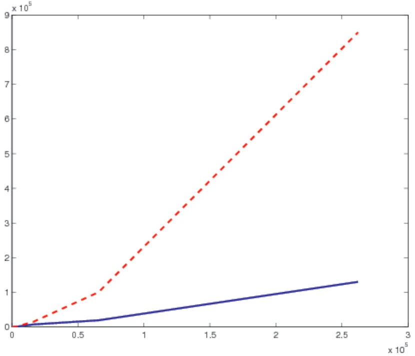
Compare the diffusion of a Dirac pulse using approximate linear solvers and exact factorization, as a function of correlation length:



Note: cost(1 MG V-cycle) $\approx$ cost(4 CG iterations)

Comparing the computational costs of CG vs MG

Consider a 2D shallow-water system and perform a complete data assimilation exercize
(3 level-1 iterations, 15 level-2 iterations, $p = 6$, tol $= 10^{-4}$)



Number of matrix-vectors products in the solution of involved linear systems as a function of problem size

Conclusions

- Use of diffusion-based correlation models in 4DVAR
- Linear algebra crucial for reasonable performance
- MG outperforms CG when use in the integration loop
- Spatial correlation properties preserved by approximate linear solves
- More arguments using eigenvalues and condition numbers

Many thanks for your attention!