Linear Least Squares for Data Assimilation

S. Gratton[1] S. Gurol[2] Ph.L. Toint[3] J. Tshimanga[1]

[1]ENSEEIHT, Toulouse, France
[2]CERFACS, Toulouse, France
[3]FUNDP, Namur, Belgium

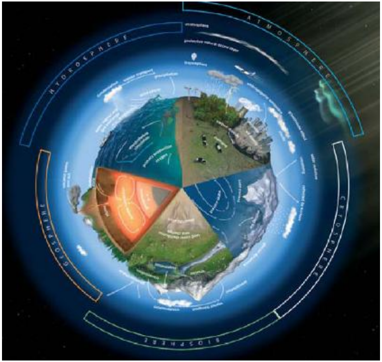
Journée de présentation des activités des étudiants

Outline

Outline

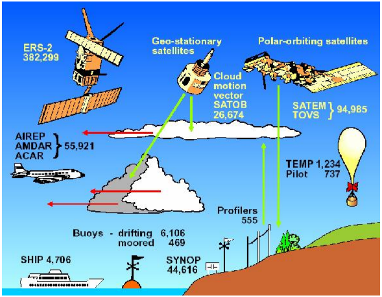
1. Introduction
2. Dual space Krylov methods
3. Conclusion and Perspectives

Data Assimilation Problems



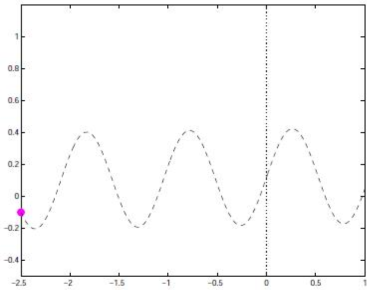
- Weather forecast
- The ocean's average temperature forecast
- Aerosol estimation
- Air quality forecast
- Soil moisture estimation
- Gravity field estimation
- ...

Data



- Type and number of observations used to estimate the atmosphere initial conditions during a typical day
- May involve up to 1,000,000,000 variables!

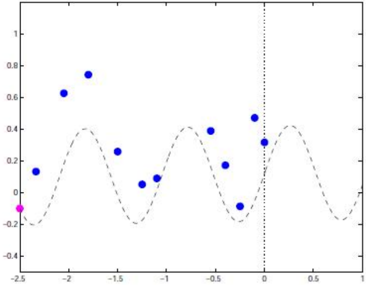
The Principle



temp. vs. days

- Known situation 2.5 days ago and background prediction

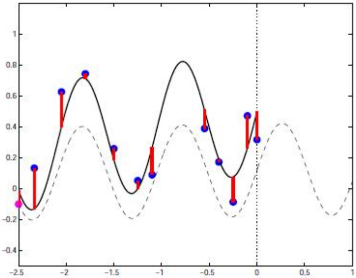
The Principle



temp. vs. days

- Known situation 2.5 days ago and background prediction
- Record temperature for the past 2.5 days

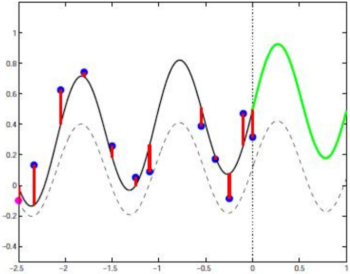
The Principle



temp. vs. days

- Known situation 2.5 days ago and background prediction
- Record temperature for the past 2.5 days
- Run the model to minimize difference between model and observations

The Principle



temp. vs. days

- Known situation 2.5 days ago and background prediction
- Record temperature for the past 2.5 days
- Run the model to minimize difference between model and observations
- Predict temperature for the next day

Four-Dimensional Variational (4D-Var) formulation

→ Very large-scale nonlinear weighted least-squares problem:

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}||x - x_b||^2_{B^{-1}} + \frac{1}{2}\sum_{j=0}^{N}||\mathcal{H}_j(\mathcal{M}_j(x)) - y_j||^2_{R_j^{-1}}$$

where:

- Size of real (operational) problems: $x, x_b \in \mathbb{R}^{10^6}$, $y_j \in \mathbb{R}^{10^5}$
- The observations y_j and the background x_b are noisy
- $\mathcal{M}_j$ are model operators (nonlinear)
- $\mathcal{H}_j$ are observation operators (nonlinear)
- B is the covariance background error matrix
- R_j are covariance observation error matrices

Incremental 4D-Var

Let rewrite the problem as:

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} ||\rho(x)||_2^2$$

Incremental 4D-Var is an inexact/truncated Gauss-Newton algorithm:

- It linearizes ρ around the current iterate $\tilde{x}$ and solves

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|\rho(\tilde{x}) + J(\tilde{x})(x - \tilde{x})\|_2^2$$

  where $J(\tilde{x})$ is the Jacobian of $\rho(x)$ at $\tilde{x}$

- It thus solves a sequence of linear systems (normal equations)

$$J^T(\tilde{x}) J(\tilde{x})(x - \tilde{x}) = -J^T(\tilde{x}) \rho(\tilde{x})$$

  where the matrix is symmetric positive definite and varies along the iterations

Incremental 4D-Var

1. Solve

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\delta\mathbf{x}_0 = \mathbf{B}^{-1}(\mathbf{x}^b - \mathbf{x}_0) + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{d}.$$

2. Perform update

$$\mathbf{x}^{(k+1)}(t_0) = \mathbf{x}^{(k)}(t_0) + \delta\mathbf{x}_0^{(k)}.$$

Exact solution writes

$$\mathbf{x}^b - \mathbf{x}_0 + \left(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{d} - \mathbf{H}(\mathbf{x}^b - \mathbf{x}_0)\right)$$

Outline

Dual formulation

Defining the constraint as $\varepsilon = \mathcal{H}(x) - y$ we can write Lagrangian function for the problem as

$$\mathcal{L}(x, \varepsilon, \lambda) = \frac{1}{2}\|x - x_b\|^2_{B^{-1}} + \frac{1}{2}\|\varepsilon\|_{R^{-1}} + \lambda^T(\varepsilon - \mathcal{H}(x) + y)$$

From Karush-Kuhn-Tucker conditions, the following stationary conditions are satisfied at any optimum:

$$\begin{aligned}
\nabla_x \mathcal{L}(x, \varepsilon, \lambda) &= B^{-1}(x - x_b) - H^T \lambda = 0 \\
\nabla_\varepsilon \mathcal{L}(x, \varepsilon, \lambda) &= R^{-1}\varepsilon = 0 \\
\nabla_\lambda \mathcal{L}(x, \varepsilon, \lambda) &= \varepsilon - \mathcal{H}(x) + y = 0
\end{aligned}$$

Dual formulation

In a matrix form:

$$\begin{bmatrix} B^{-1} & 0 & -H^T \\ 0 & R^{-1} & I \\ -H & I & 0 \end{bmatrix} \begin{bmatrix} x \\ \varepsilon \\ \lambda \end{bmatrix} = \begin{bmatrix} B^{-1}x_b \\ 0 \\ -y \end{bmatrix},$$

after elimination:

$$\begin{bmatrix} I & 0 & -BH^T \\ 0 & I & R \\ 0 & I & -HBH^T \end{bmatrix} \begin{bmatrix} x \\ \varepsilon \\ \lambda \end{bmatrix} = \begin{bmatrix} x_b \\ 0 \\ -y + Hx_b \end{bmatrix}$$

Dual formulation

Therefore, the solution can be written as

$$\lambda = (HBH^T + R)^{-1}(y - Hx_b)$$
$$x = x_b + BH^T\lambda$$
$$\varepsilon = -R\lambda$$

We therefore obtain

$$x = x_b + BH^T(HBH^T + R)^{-1}(y - Hx_b)$$

which can be obtained also from the solution in primal space using Sherman-Morrison formula.

Dual formulation

1. Iteratively solve

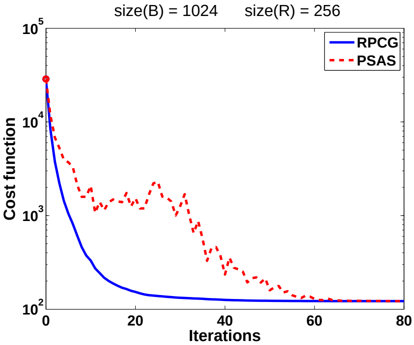
$$\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\right)\lambda = \mathbf{y} - \mathbf{H}\mathbf{x}_b \quad \text{for} \quad \lambda$$

2. Set

$$\delta x_0 = \mathbf{x}_b - \mathbf{x}_0 + \mathbf{B}\mathbf{H}^{\mathrm{T}}\lambda$$

- Very popular when few observations compared to model variables.
- PSAS algorithm (Courtier 1997): PCG with canonical inner product
- RPCG algorithm (Gratton and Tschimanga 2009): PCG with $\mathbf{HBH^T}$ inner product

Experiments



CG-like algorithm : assumptions 1

1. Suppose the CG algorithm is applied to solve the Inc-4D using a preconditioning matrix $\mathbf{F}$
2. Suppose there exists $\mathbf{G}^{m\times m}$ such that

$$\mathbf{F}\mathbf{H}^{\mathrm{T}} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{G}$$

3. For "exact" preconditioners

$$\left(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{I} + \mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\right)^{-1}$$

Preconditioned CG on Incremental 4D-Var cost function

Initialization steps

Loop: WHILE

1. $\mathbf{q}_{i-1} = \mathbf{A}\mathbf{p}_{i-1}$
2. $\alpha_{i-1} = \mathbf{r}_{i-1}^{\mathrm{T}}\mathbf{z}_{i-1} \,/\, \mathbf{q}_{i-1}^{\mathrm{T}}\mathbf{p}_{i-1}$
3. $\mathbf{v}_i = \mathbf{v}_{i-1} + \alpha_{i-1}\mathbf{p}_{i-1}$
4. $\mathbf{r}_i = \mathbf{r}_{i-1} + \alpha_{i-1}\mathbf{q}_{i-1}$
5. $\mathbf{z}_i = \mathbf{F}\mathbf{r}_i$
6. $\beta_i = \mathbf{r}_i^{\mathrm{T}}\mathbf{z}_i \,/\, \mathbf{r}_{i-1}^{\mathrm{T}}\mathbf{z}_{i-1}$
7. $\mathbf{p}_i = -\mathbf{z}_i + \beta_i\mathbf{p}_{i-1}$

Initialization steps

Loop: WHILE

1. $\mathbf{q}_{i-1} = (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})\mathbf{p}_{i-1}$
2. $\alpha_{i-1} = \mathbf{r}_{i-1}^{\mathrm{T}}\mathbf{z}_{i-1} \,/\, \mathbf{q}_{i-1}^{\mathrm{T}}\mathbf{p}_{i-1}$
3. $\mathbf{v}_i = \mathbf{v}_{i-1} + \alpha_{i-1}\mathbf{p}_{i-1}$
4. $\mathbf{r}_i = \mathbf{r}_{i-1} + \alpha_{i-1}\mathbf{q}_{i-1}$
5. $\mathbf{z}_i = \mathbf{F}\mathbf{r}_i$
6. $\beta_i = \mathbf{r}_i^{\mathrm{T}}\mathbf{z}_i \,/\, \mathbf{r}_{i-1}^{\mathrm{T}}\mathbf{z}_{i-1}$
7. $\mathbf{p}_i = -\mathbf{z}_i + \beta_i\mathbf{p}_{i-1}$

An useful observation

Theorem

Suppose that

1. $\mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{G} = \mathbf{F}\mathbf{H}^{\mathrm{T}}$.
2. $\mathbf{v}_0 = \mathbf{x}^b - \mathbf{x}_0$.

$\rightarrow$ *vectors* $\widehat{\mathbf{r}}_i$, $\widehat{\mathbf{p}}_i$, $\widehat{\mathbf{v}}_i$, $\widehat{\mathbf{z}}_i$ *and* $\widehat{\mathbf{q}}_i$ *such that*

$$\begin{aligned} \mathbf{r}_i &= \mathbf{H}^{\mathrm{T}}\widehat{\mathbf{r}}_i, \\ \mathbf{p}_i &= \mathbf{B}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{p}}_i, \\ \mathbf{v}_i &= \mathbf{v}_0 + \mathbf{B}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{v}}_i, \\ \mathbf{z}_i &= \mathbf{B}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{z}}_i, \\ \mathbf{q}_i &= \mathbf{H}^{\mathrm{T}}\widehat{\mathbf{q}}_i \end{aligned}$$

Restricted PCG (version 1)

Initialization steps

given $\mathbf{v}_0$; $\mathbf{r}_0 = (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})\mathbf{v}_0 - \mathbf{b}$, ...

Loop: WHILE

1. $\widehat{\mathbf{q}}_{i-1} = (\mathbf{I}_m + \mathbf{R}^{-1}\mathbf{H}\mathbf{B}^{-1}\mathbf{H}^{\mathrm{T}})\widehat{\mathbf{p}}_{i-1}$
2. $\alpha_{i-1} = \widehat{\mathbf{r}}_{i-1}^{\mathrm{T}}\mathbf{H}\mathbf{B}\mathbf{H}^{T}\widehat{\mathbf{z}}_{i-1} \,/\, \widehat{\mathbf{q}}_{i-1}^{\mathrm{T}}\mathbf{H}\mathbf{B}\mathbf{H}^{T}\widehat{\mathbf{p}}_{i-1}$
3. $\widehat{\mathbf{v}}_i = \widehat{\mathbf{v}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{p}}_{i-1}$
4. $\widehat{\mathbf{r}}_i = \widehat{\mathbf{r}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{q}}_{i-1}$
5. $\widehat{\mathbf{z}}_i = \mathbf{F}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{r}}_i = \mathbf{G}\widehat{\mathbf{r}}_i$
6. $\beta_i = \widehat{\mathbf{r}}_i^{\mathrm{T}}\mathbf{H}\mathbf{B}\mathbf{H}^{T}\widehat{\mathbf{z}}_i \,/\, \widehat{\mathbf{r}}_{i-1}^{\mathrm{T}}\mathbf{H}\mathbf{B}\mathbf{H}^{T}\widehat{\mathbf{z}}_{i-1}$
7. $\widehat{\mathbf{p}}_i = -\widehat{\mathbf{z}}_i + \beta_i\widehat{\mathbf{p}}_{i-1}$

More transformations

1. Consider $\mathbf{w}$ and $\mathbf{t}$ defined by

$$\mathbf{w}_i = \mathbf{HBH}^{\mathrm{T}}\widehat{\mathbf{z}}_i \quad \text{and} \quad \mathbf{t}_i = \mathbf{HBH}^{\mathrm{T}}\widehat{\mathbf{p}}_i$$

2. From Restricted PCG (version 1)

$$\mathbf{t}_i = \begin{cases} -\mathbf{w}_0 & \text{if } \ i = 0, \\ -\mathbf{w}_i + \beta_i \mathbf{t}_{i-1} & \text{if } \ i > 0, \end{cases}$$

3. Use these relations into Restricted PCG (version 1)
4. Transform Restricted PCG (version 1) into Restricted PCG (version 2)

Restricted PCG (version 2)

Initialization steps

Loop: WHILE

1. $\widehat{\mathbf{q}}_{i-1} = \mathbf{R}^{-1}\mathbf{t}_{i-1} + \widehat{\mathbf{p}}_{i-1}$
2. $\alpha_{i-1} = \mathbf{w}_{i-1}^{\mathrm{T}}\widehat{\mathbf{r}}_{i-1} \,/\, \widehat{\mathbf{q}}_{i-1}^{\mathrm{T}}\mathbf{t}_{i-1}$
3. $\widehat{\mathbf{v}}_{i} = \widehat{\mathbf{v}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{p}}_{i-1}$
4. $\widehat{\mathbf{r}}_{i} = \widehat{\mathbf{r}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{q}}_{i-1}$
5. $\widehat{\mathbf{z}}_{i} = \mathbf{G}\widehat{\mathbf{r}}_{i}$
6. $\mathbf{w}_{i} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{z}}_{i}$
7. $\beta_{i} = \mathbf{w}_{i}^{\mathrm{T}}\widehat{\mathbf{r}}_{i} \,/\, \mathbf{w}_{i-1}^{\mathrm{T}}\widehat{\mathbf{r}}_{i-1}$
8. $\widehat{\mathbf{p}}_{i} = -\widehat{\mathbf{z}}_{i} + \beta_{i}\widehat{\mathbf{p}}_{i-1}$
9. $\mathbf{t}_{i} = -\mathbf{w}_{i} + \beta_{i}\mathbf{t}_{i-1}$

$\mathbf{F}$ as a Quasi-Newton Limited Memory Preconditioner

- Quasi-Newton LMPs are simply based on the idea that generates preconditioners by using LBFGS updating formula

$$\mathbf{F_{k+1}} = (\mathbf{I_n} - \tau_\mathbf{k}\mathbf{p_k}\mathbf{q_k^T})\mathbf{F_k}(\mathbf{I_n} - \tau_\mathbf{k}\mathbf{q_k}\mathbf{p_k^T}) + \tau_\mathbf{k}\mathbf{p_k}\mathbf{p_k^T}$$

- $\tau_\mathbf{k} = \mathbf{1}/(\mathbf{q_k^T}\mathbf{p_k})$
- $\mathbf{q_k} = (\mathbf{B^{-1}} + \mathbf{H^T}\mathbf{R^{-1}}\mathbf{H})\mathbf{p_k}$

- $\Delta\mathbf{F_k}$ defined by $\Delta\mathbf{F_k} = \mathbf{F_{k+1}} - \mathbf{F_k}$, is the optimal solution to the problem:

$$\min_{\Delta\mathbf{F_k}} \left\| \mathbf{W^{1/2}}\boldsymbol{\Delta}\mathbf{F_k}\mathbf{W^{1/2}} \right\|_F$$

$$\text{subject to} \quad \Delta\mathbf{F_k} = \boldsymbol{\Delta}\mathbf{F_k^T}, \quad \mathbf{F_{k+1}}\mathbf{q_k} = \mathbf{p_k}$$

where $\mathbf{W}$ is any symmetric positive definite matrix satisfying $\mathbf{W}\mathbf{p_k} = \mathbf{q_k}$

$\mathbf{G}$ as a Quasi-Newton Limited Memory Preconditioner

- Giving $\mathbf{F}$, we can find $\mathbf{G}$ that satisfies $\mathbf{F}\mathbf{H}^{\mathbf{T}} = \mathbf{B}\mathbf{H}^{\mathbf{T}}\mathbf{G}$
- $\mathbf{G}$ can be derived by using the formula for $\mathbf{F}$ using the relations

$$\begin{aligned} \mathbf{p}_i &= \mathbf{B}\mathbf{H}^{\mathrm{T}}\widehat{\mathbf{p}}_i, \\ \mathbf{q}_i &= \mathbf{H}^{\mathrm{T}}\widehat{\mathbf{q}}_i \end{aligned}$$

$$\mathbf{G_{k+1}} = (\mathbf{I_m} - \widehat{\tau}_\mathbf{k}\widehat{\mathbf{p}}_\mathbf{k}(\mathbf{M}\widehat{\mathbf{q}}_\mathbf{k})^{\mathbf{T}})\mathbf{G_k}(\mathbf{I_m} - \widehat{\tau}_\mathbf{k}\widehat{\mathbf{q}}_\mathbf{k}\widehat{\mathbf{p}}_\mathbf{k}^{\mathbf{T}}\mathbf{M}) + \widehat{\tau}_\mathbf{k}\widehat{\mathbf{p}}_\mathbf{k}\widehat{\mathbf{p}}_\mathbf{k}^{\mathbf{T}}\mathbf{M}$$

- $\mathbf{M} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}}$
- $\widehat{\mathbf{p}}_\mathbf{k}$ is the search direction
- $\widehat{\mathbf{q}}_\mathbf{k} = (\mathbf{I_m} + \mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}})\widehat{\mathbf{p}}_\mathbf{k}$
- $\widehat{\tau}_\mathbf{k} = 1/(\widehat{\mathbf{q}}_\mathbf{k}^{\mathbf{T}}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}}\widehat{\mathbf{p}}_\mathbf{k})$

$\mathbf{G}$ as a Quasi-Newton Limited Memory Preconditioner

$\Delta\mathbf{G_k}$ defined by $\Delta\mathbf{G_k} = \mathbf{G_{k+1}} - \mathbf{G_k}$ is the optimum solution to the minimization problem defined as:

$$\min_{\mathbf{\Delta G_k}} \left\| W^{1/2} M^{1/2} \Delta G_k M^{-1/2} W^{1/2} \right\|_F$$

$$\text{subject to} \quad \mathbf{M\Delta G_k} = \mathbf{\Delta G_k^T M}, \quad \mathbf{G_{k+1}}\widehat{\mathbf{q}}_\mathbf{k} = \widehat{\mathbf{p}}_\mathbf{k}$$

The norm in this problem is considered as a weighted Frobenius norm, where W is any symmetric positive definite matrix satisfying $\mathbf{W M^{1/2}}\widehat{\mathbf{p}}_\mathbf{k} = \mathbf{M^{1/2}}\widehat{\mathbf{q}}_\mathbf{k}$ and $\mathbf{M} = \mathbf{HBH^T}$.

Restricted PCG (version 3)

1. Consider a new vector $\mathbf{l}$ is defined as

$$\mathbf{l_i} = \mathbf{HBH^T\widehat{r}_i}$$

2. $\mathbf{\widehat{z}_i} = \mathbf{G\widehat{r}_i}$ and $\mathbf{w_i} = \mathbf{HBH}^{\mathrm{T}}\mathbf{\widehat{z}_i}$
3. $\mathbf{HBH^TG}$ is symmetric $(\mathbf{HFH^T} = \mathbf{HBH^TG})$

$$\mathbf{w_i} = \mathbf{HBH^TG\widehat{r}_i} = \mathbf{G^THBH^T\widehat{r}_i} = \mathbf{G^Tl_i}$$

4. Multiply line 4 of Restricted PCG (version 2) with $\mathbf{HBH^T}$ gives

$$\mathbf{HBH^T\widehat{q}_i} = (\mathbf{l_i} - \mathbf{l_{i-1}})/\alpha_\mathbf{i}$$

5. Use these relations into Restricted PCG (version 2)
6. Transform Restricted PCG (version 2) into Restricted PCG (version 3)

Restricted PCG (version 3)

Loop: WHILE

1. $\widehat{\mathbf{q}}_{i-1} = \mathbf{R}^{-1}\mathbf{t}_{i-1} + \widehat{\mathbf{p}}_{i-1}$
2. $\alpha_{i-1} = \mathbf{w}_{i-1}^{\mathrm{T}}\widehat{\mathbf{r}}_{i-1} \,/\, \widehat{\mathbf{q}}_{i-1}^{\mathrm{T}}\mathbf{t}_{i-1}$
3. $\widehat{\mathbf{v}}_i = \widehat{\mathbf{v}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{p}}_{i-1}$
4. $\widehat{\mathbf{r}}_i = \widehat{\mathbf{r}}_{i-1} + \alpha_{i-1}\widehat{\mathbf{q}}_{i-1}$
5. $\mathbf{l}_i = \mathbf{HBH}^{\mathrm{T}}\widehat{\mathbf{r}}_i$
6. $\widehat{\mathbf{z}}_i = \mathbf{G}\widehat{\mathbf{r}}_i$
7. $\mathbf{w}_i = \mathbf{G}^{\mathrm{T}}\mathbf{l}_i$
8. $\beta_i = \mathbf{w}_i^{\mathrm{T}}\widehat{\mathbf{r}}_i \,/\, \mathbf{w}_{i-1}^{\mathrm{T}}\widehat{\mathbf{r}}_{i-1}$
9. $\widehat{\mathbf{p}}_i = -\widehat{\mathbf{z}}_i + \beta_i\widehat{\mathbf{p}}_{i-1}$
10. $\mathbf{t}_i = -\mathbf{w}_i + \beta_i\mathbf{t}_{i-1}$
11. $\mathbf{mq}_i = \mathbf{l}_i - \mathbf{l}_{i-1} \,/\, \alpha_i$
12. Calculate G

Convergence Properties

- If FA has eigenvalues $\mu_1 \leq \mu_2 \leq ... \leq \mu_n$, PCG algorithm with zero initial starting vector satisfies the inequality:

$$\|x_{k+1} - x^*\|_A \leq 2(\frac{\sqrt{\mu_n} - \sqrt{\mu_1}}{\sqrt{\mu_n} + \sqrt{\mu_1}})^k \|x^*\|_A$$

- When RPCG is used, the iterates are belongs to the affine subspace of $x_0 + Im(BH^T)$.
- If $\mathbf{G}\widehat{\mathbf{A}}$ has eigenvalues $\nu_\mathbf{1} \leq \nu_2 \leq ... \leq \nu_m$, Restricted PCG (version 3) with zero initial starting vector satisfies the inequality:

$$\|\mathbf{x_{k+1}} - \mathbf{x}^*\|_\mathbf{A} \leq \mathbf{2}(\frac{\sqrt{\nu_\mathbf{m}} - \sqrt{\nu_\mathbf{1}}}{\sqrt{\nu_\mathbf{m}} + \sqrt{\nu_\mathbf{1}}})^\mathbf{k} \|\mathbf{x}^*\|_\mathbf{A}$$

where $\widehat{\mathbf{A}} = \mathbf{I} + \mathbf{R}^{-1}\mathbf{HBH^T}$

Convergence Properties

- $FABH^T = BH^T G\widehat{A}$. Therefore; BH^T is an invariant subspace of FA
- Every eigenvalue of $G\widehat{A}$ is an eigenvalue of FA. So, $\mu_1 \leq \nu_1$ and $\mu_n \geqslant \nu_n$.

If $\mathbf{FA}$ has eigenvalues $\mu_1 \leq \mu_2 \leq ... \leq \mu_n$ and $\mathbf{G\widehat{A}}$ has eigenvalues $\nu_\mathbf{1} \leq \nu_2 \leq ... \leq \nu_m$, Restricted PCG (version 3) with zero initial starting vector satisfies the inequality:

$$\|x_{k+1} - x^*\|_A \leq 2\left(\frac{\sqrt{\nu_m} - \sqrt{\nu_1}}{\sqrt{\nu_m} + \sqrt{\nu_1}}\right)^k \|x^*\|_A \leq 2\left(\frac{\sqrt{\mu_n} - \sqrt{\mu_1}}{\sqrt{\mu_n} + \sqrt{\mu_1}}\right)^k \|x^*\|_A$$

Loss (and recovery) of orthogonality

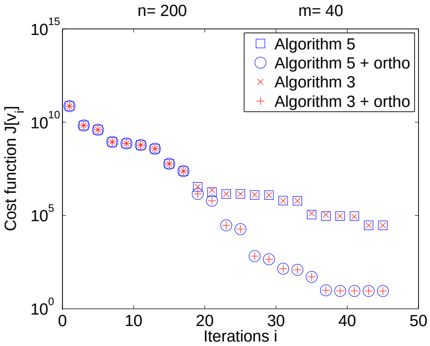
1. The modified (G-S) orthogonalization scheme writes

$$\mathbf{r}_i \leftarrow \prod_{j=1}^{i-1} \left(\mathbf{I}_n - \frac{\mathbf{r}_j \mathbf{r}_j^{\mathrm{T}}}{\mathbf{r}_j^{\mathrm{T}} \mathbf{F} \mathbf{r}_j} \right) \mathbf{r}_i .$$

2. We suggest the following re-orthogonalization scheme

$$\hat{\mathbf{r}}_i \leftarrow \prod_{j=1}^{i-1} \left(\mathbf{I}_m - \frac{\hat{\mathbf{r}}_j \mathbf{w}_j^{\mathrm{T}}}{\hat{\mathbf{r}}_j^{\mathrm{T}} \mathbf{w}_j} \right) \hat{\mathbf{r}}_i .$$

Loss (and recovery) of orthogonality : experiment



Outline

Conclusions

- Explain PCG methods in dual space
- RPCG is mathematically equivalent to PCG in the sense that, in exact arithmetic, both algorithms generate exactly the same iterates.
- Cheaper than CG (memory and computation)
- Possible to find $\mathbf{G}$ that satisfies $\mathbf{F}\mathbf{H}^{\mathbf{T}} = \mathbf{B}\mathbf{H}^{\mathbf{T}}\mathbf{G}$ for a given $\mathbf{F}$
- Some numerical experiments shown

Perpectives

Perpectives

- Behaviour in presence of round-off error
- Find efficient preconditioners $\mathbf{F}$ such that

$$\mathbf{F}\mathbf{H}^{\mathrm{T}} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{G}$$

- Implement RPCG in a real life data assimilation system : RTRA project

Thank you for your attention !