On the use of saddle formulation in weakly-constrained 4D-Var

Serge Gratton*, Selime Gürol†, Ehouarn Simon* and Philippe L. Toint‡

* INPT, IRIT, Toulouse, France
† CERFACS, Toulouse, France
‡ NAXYS, University of Namur, Belgium
ehouarn.simon@enseeiht.fr

2 July 2018

Outline

Outline

Weakly-constrained 4D-Var

A large-scale weighted nonlinear least-squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} J(\mathbf{x}) = \frac{1}{2}\|x_0 - x_b\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\sum_{j=0}^{N_{sw}} \|\mathcal{H}_j(x_j) - y_j\|^2_{\mathbf{R}_j^{-1}} + \frac{1}{2}\sum_{j=1}^{N_{sw}} \|x_j - \mathcal{M}_j(x_{j-1})\|^2_{\mathbf{Q}_j^{-1}}$$

where

- $\mathbf{x} = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N_{sw}})^T \in \mathbb{R}^n$ is the control variable (with $x_j = x(t_j)$),
- $\mathbf{x}_b$ is the background given at the initial time (t_0),
- $\mathbf{y}_j \in \mathbb{R}^{m_j}$ is the observation vector over a given time interval,
- $\mathcal{H}_j$ maps the state vector $\mathbf{x}_j$ from model space to observation space,
- $\mathcal{M}_j$ represents an integration of the numerical model from time t_{j-1} to t_j,
- $\mathbf{B}$, $\mathbf{R}_j$ and $\mathbf{Q}_j$ are the covariances of the background, observation and model error.

▷ Model error, longer time windows, accumulation of more observations, but larger problems.

Minimization: truncated Gauss-Newton method

- Linearizing $\mathcal{M}$ and $\mathcal{H}$ at the current iterate.
- Minimizing the resulting quadratic function.

The linearized subproblems

Outer iteration k

$$\min_{\delta x \in \mathbb{R}^n} q_{st}(\delta \mathbf{x}) = \frac{1}{2}\|\mathbf{L}\delta\mathbf{x} - \mathbf{b}\|^2_{\mathbf{D}^{-1}} + \frac{1}{2}\|\mathbf{H}\delta\mathbf{x} - \mathbf{d}\|^2_{\mathbf{R}^{-1}}$$

where

- $$\mathbf{L} = \begin{pmatrix} \mathbf{I}_n & & & & \\ -\mathbf{M}_1^{(k)} & \mathbf{I}_n & & & \\ & -\mathbf{M}_2^{(k)} & \mathbf{I}_n & & \\ & & \ddots & \ddots & \\ & & & -\mathbf{M}_{N_{sw}}^{(k)} & \mathbf{I}_n \end{pmatrix},$$

  ▷ Parallel-in-time matrix-vector products.

- $$\mathbf{d} = \begin{pmatrix} y_0 - \mathcal{H}_0(x_0^{(k)}) \\ y_1 - \mathcal{H}_1(x_1^{(k)}) \\ \vdots \\ y_{N_{sw}} - \mathcal{H}_{N_{sw}}(x_{N_{sw}}^{(k)}) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} x_0^{(k)} - x_b \\ \mathcal{M}_1(x_0^{(k)}) - x_1^{(k)} \\ \vdots \\ \mathcal{M}_{N_{sw}}(x_{N_{sw}-1}^{(k)}) - x_{N_{sw}}^{(k)} \end{pmatrix},$$

- $\mathbf{H} = \mathrm{diag}(\mathbf{H}_0^{(k)}, \mathbf{H}_1^{(k)}, \dots, \mathbf{H}_{N_{sw}}^{(k)})$,
- $\mathbf{D} = \mathrm{diag}(\mathbf{B}, \mathbf{Q}_1, \dots, \mathbf{Q}_{N_{sw}})$ and $\mathbf{R} = \mathrm{diag}(\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_{N_{sw}})$.

Saddle Point Approach

Reformulation of the problem

- Let us consider weak-constraint 4D-Var as a constrained problem:

$$\min_{(\boldsymbol{\delta}\mathbf{p},\boldsymbol{\delta}\mathbf{w},\boldsymbol{\delta}\mathbf{x})} \frac{1}{2}\|\boldsymbol{\delta}\mathbf{p}-\mathbf{b}\|^2_{\mathbf{D}^{-1}} + \frac{1}{2}\|\boldsymbol{\delta}\mathbf{w}-\mathbf{d}\|^2_{\mathbf{R}^{-1}}$$

$$\text{subject to} \quad \boldsymbol{\delta}\mathbf{p} = \mathbf{L}\boldsymbol{\delta}\mathbf{x} \quad \text{and} \quad \delta\mathbf{w} = \mathbf{H}\boldsymbol{\delta}\mathbf{x}$$

- The *Lagrangian function* for this problem reads

$$\begin{aligned}\mathcal{L}(\delta\mathbf{w},\delta\mathbf{p},\delta\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) =& \frac{1}{2}\|\boldsymbol{\delta}\mathbf{p}-\mathbf{b}\|^2_{\mathbf{D}^{-1}} + \frac{1}{2}\|\boldsymbol{\delta}\mathbf{w}-\mathbf{d}\|^2_{\mathbf{R}^{-1}} \\ &+ \boldsymbol{\lambda}^T(\delta\mathbf{p}-\mathbf{L}\delta\mathbf{x}) + \boldsymbol{\mu}^T(\delta\mathbf{w}-\mathbf{H}\delta\mathbf{x})\end{aligned}$$

- The stationnary point of $\mathcal{L}$ satisfies the following equations:

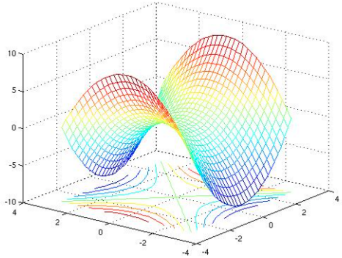
$$\begin{aligned}\mathbf{D}^{-1}(\mathbf{L}\delta\mathbf{x}-\mathbf{b}) + \boldsymbol{\lambda} &= 0 \\ \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x}-\mathbf{d}) + \boldsymbol{\mu} &= 0 \\ \mathbf{L}^T\boldsymbol{\lambda} + \mathbf{H}^T\boldsymbol{\mu} &= 0\end{aligned}$$

Saddle Point Approach

- In matrix form:

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

- The solution of this problem is a saddle point, with no inverse of covariance matrix involved.



▷ Solution algorithm: iterative methods (MINRES, GMRES, ...) with a preconditioner.

The original saddle method: M. Fisher

Consider the solution of $r(\delta\lambda, \delta\mu, \delta x) = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{L} \\ \mathbf{0} & \mathbf{R} & \mathbf{H} \\ \mathbf{L}^{\mathrm{T}} & \mathbf{H}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \delta\lambda \\ \delta\mu \\ \delta x \end{pmatrix} - \begin{pmatrix} \mathbf{b} \\ \mathbf{d} \\ \mathbf{0} \end{pmatrix} = \mathbf{0}$

Saddle-original (SAQ0)

While not converged:

1. **Compute** $J(x_k)$ and $g_k = \nabla_x J(x_k)$
2. **Apply** the preconditioned GMRES algorithm to solve the system $r(\delta\lambda, \delta\mu, \delta x) = 0$. Terminate the iterations if $\|r(\delta\lambda, \delta\mu, \delta x)\| \leq \varepsilon_r(\|b\| + \|d\|)$ or $j = n_{inner}$ to yield δx_k
3. **Set** $x_{k+1} = x_k + \delta x_k$

Possible preconditioners

$$P_M = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \tilde{\mathbf{L}} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \\ \tilde{\mathbf{L}}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad P_B = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} \end{pmatrix}, \quad P_T = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \tilde{\mathbf{L}} \\ \mathbf{0} & \mathbf{R} & \mathbf{H} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} \end{pmatrix}$$

with $\mathbf{S}^{-1} = \tilde{\mathbf{L}}^{-1}\mathbf{D}\tilde{\mathbf{L}}^{-\mathbf{T}}$ and $\tilde{\mathbf{L}} \sim \mathbf{L}$ (square, nonsingular),

Outline

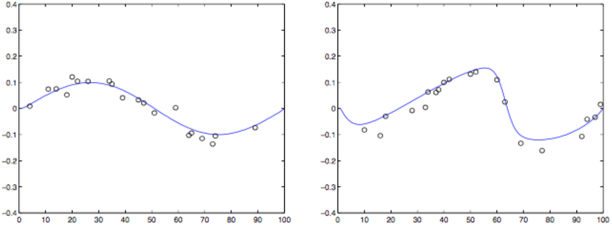
Numerical experiments

A Burgers system

- We consider the one dimensional dynamical system

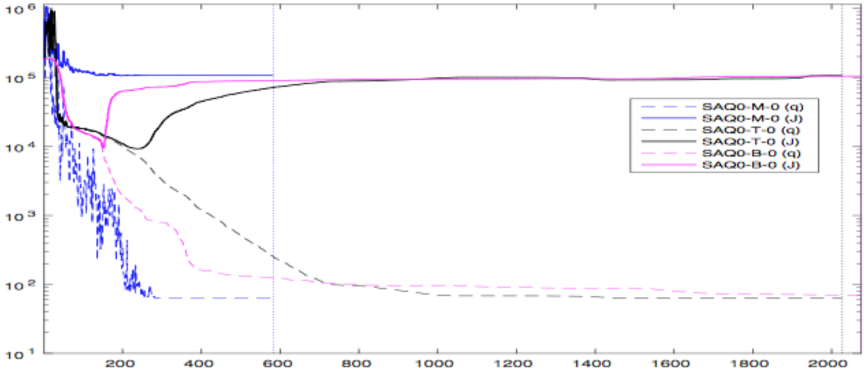
$$\begin{cases} \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} - \nu\dfrac{\partial^2 u}{\partial x^2} = f(x) \\ \qquad (x,t) \in]0,1[\times\mathbb{R}_+^* \\ u(0,t) = u(1,t) = 0, \quad t > 0 \\ u(x,0) = k\sin(\pi x)\sin(\pi(1-x)); \\ \qquad x \in]0,1[\end{cases}$$

- The field u is partially observed in space and time.



Reference trajectory and observations at the end of the first and last subwindow ($N_{SW} = 50$).

The original saddle method

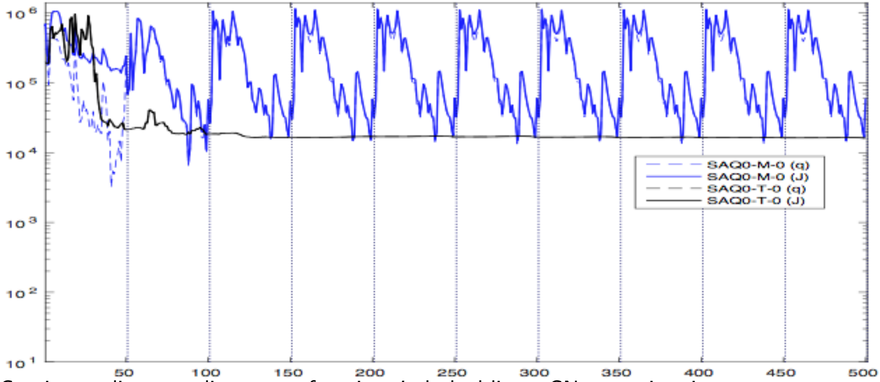


Continuous lines: nonlinear cost function J; dashed lines: GN approximation q_{st}.

One outer iteration: fully accurate ($\varepsilon_r = 10^{-12}$).

- Nonlinear problem: the values of J and q_{st} do not agree for large steps $\delta\mathbf{x}$.
- No convergence in one outer iteration (optimal value ~ 63.11).
- Non-monotonic evolutions of q_{st} and J (obvious with the preconditioner P_M).

The original saddle method



Continuous lines: nonlinear cost function J; dashed lines: GN approximation q_{st}.

10 outer iterations: 50 inner iterations ($\varepsilon_r = 10^{-7}$).

- Good fit between q_{st} and J beyond the first iteration for moderately small steps.
- No significant reduction in J (optimal value ~ 63.11).
- Non-monotonic evolutions of q_{st} and J.
  - ▷ Stoppping cannot solely rely on maximum number of iterations.

Outline

A safeguarded saddle algorithm

Safeguarded Saddle (SAQℓ)

While not converged:

1. **Compute** $J(x_k)$ and $g_k = \nabla_x J(x_k)$
2. **Apply** the preconditioned GMRES algorithm to solve the system $r(\delta\lambda, \delta\mu, \delta x) = 0$. At iteration j, terminate if

$$\text{mod }(j, \ell) = 0 \quad \text{and} \quad q_{st}(0) - q_{st}(\delta \mathbf{x}) \geq \max\left[\epsilon_q \min[1, \|\mathbf{g}_k\|^2], \theta_j\right]$$

   or if the residual equation is solved to full accuracy, yielding a step δx_k.
3. Perform a backtracking linesearch on J along $\delta \mathbf{x}_k$ yielding a step-length α_k.
4. **Set** $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \delta \mathbf{x}_k$.

with $\ell \in \mathbb{N}$ the model check frequency and $\{\theta_j\}$ a sequence that goes to zero.

A safeguarded saddle algorithm

Global convergence

- The stopping criterion and the strict convexity of q_{st} ensure that $\delta \mathbf{x}_k$ is "gradient-related":

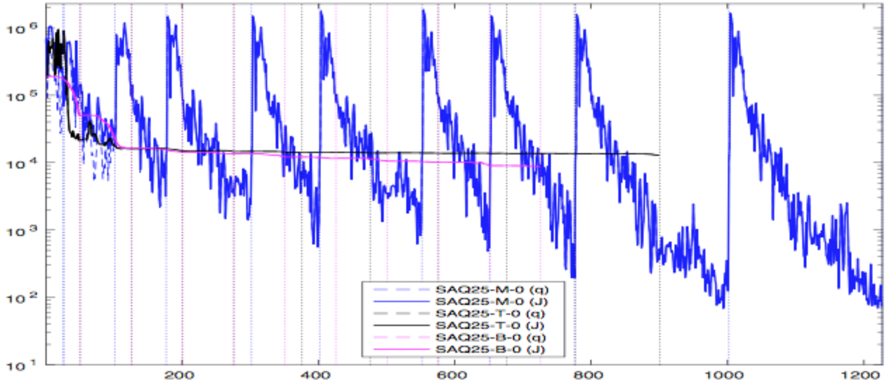
$$-g_k^T \delta x_k \geq \kappa_1 \|g_k\|^2 \text{ and } \|\delta x_k\| \leq \kappa_2 \|g_k\|$$

  for some positive constants $0 < \kappa_1 \leq \kappa_2$.
- Applying the linesearch to "gradient-related" directions is sufficient to ensure a monotonic decrease of the sequence $\{J(\mathbf{x}_k)\}$.

Remarks

- The sequence $\{\theta_j\}$ goes to zero and forces GMRES not to stop prematurely.
- The stopping criterion involves the computation of the quadratic.
  - ▷ Application of the operators $\mathbf{L}$, $\mathbf{D}^{-1}$, $\mathbf{H}$, $\mathbf{R}^{-1}$.
- The GMRES algorithm may need more iterations than previsously.
  - ▷ Potentially more expensive.

Model check frequency: every $\ell = 25$ inner iteration



Continuous lines: nonlinear cost function J; dashed lines: GN approximation q_{st}.

10 outer iterations ($\epsilon_q = 10^{-2}$)

- Better decrease of q_{st}, but more inner iterations (2 times).
- Preconditioner $\mathbf{P}_M$: better decrease in J.
- Poor performances of the preconditioners $\mathbf{P}_B$ and $\mathbf{P}_T$.

Outline

Conclusion and perspectives

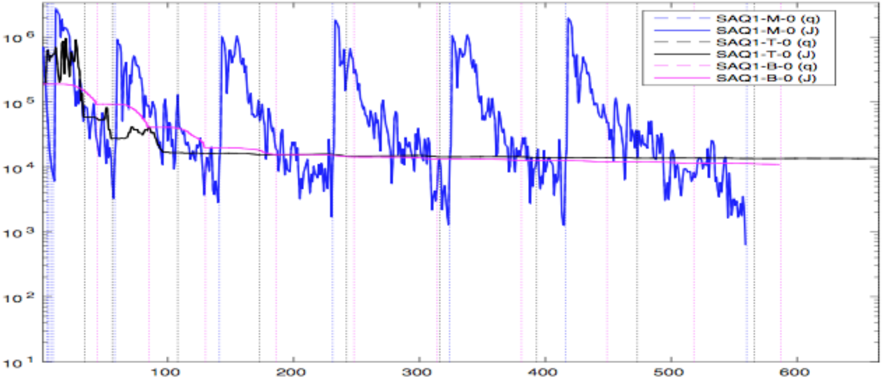
- Problematic behaviour of the original saddle formulation as general method for solving the weakly-constrained 4DVar problem.
  - ▷ Poor correlation between quadratic model decrease and reduction of the residual of the associated optimality conditions.
  - ▷ The original saddle formulation may produce reasonable results on some favourable examples (QG problem, not shown).
- A safeguarded approach focusing on quadratic reduction has been suggested.
  - ▷ Stopping criterion based on periodical evaluations of the quadratic function and use of a linesearch.
  - ▷ Decrease in the cost function values at each outer iteration despite possible chaotic behaviour in the inner iterations.
  - ▷ Increase in the computational costs (more inner iterations, operator $\mathbf{D}^{-1}$ in the quadratic function).
- Avoiding the use of $\mathbf{D}^{-1}$?
  - ▷ Approximately solving the linear system ($\mathbf{D}\mathbf{x} = \mathbf{z}$).
  - ▷ Convergence analysis for noisy/inexact computation of $\mathbf{D}^{-1}$
- Preconditioners for the saddle formulation.

Thank you!

Bibliography

- M. Fisher, S. Gratton, S. Gürol, Y. Trémolet, X. Vasseur: Low rank updates in preconditioning the saddle point systems arising from data assimilation problems. *Optimization Methods and Software*, 33(1), 45-69, 2017.
- M. Fisher and S. Gürol: Parallelisation in the time dimension of four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(703):1136-1147, 2017.
- M. Fisher, Y. Trémolet, H. Auvinen, D. Tan, and P. Poli: Weak-constrained and long window 4D-Var. Technical Report 655, ECMWF, 2011.
- S. Gratton, S. Gürol, E. Simon and P.L. Toint: A note on preconditioning weighted linear least-squares, with consequences for weakly-constrained variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, in press.
- S. Gratton, S. Gürol, E. Simon and P.L. Toint: Guaranteeing the convergence of the saddle formulation for weakly-constrained 4D-VAR data assimilation, *Quarterly Journal of the Royal Meteorological Society*, in press.
- Y. Trémolet, Accounting for an imperfect model in 4D-Var: *Quarterly Journal of the Royal Meteorological Society*, 132(621):2483-2504, 2006.
- Y. Trémolet, Model error estimation in 4D-Var: *Quarterly Journal of the Royal Meteorological Society*, 133(626):1267-1280, 2006.
- P.A. Vidard, A. Piacentini, and F.-X. Le Dimet: Variational data analysis with control of the forecast bias. *Tellus*, 56A:177-188, 2004.
- D. Zupanski: A general weak constraint applicable to operational 4DVAR data assimilation systems. *Monthly Weather Review*, 125, 1997.

Model check frequency: every $\ell = 1$ inner iteration



Continuous lines: nonlinear cost function J; dashed lines: GN approximation q_{st}.

10 outer iterations ($\epsilon_q = 10^{-2}$)

- Reduction in J every outer iteration.
- Preconditioner $\mathbf{P}_M$: oscillations but best decrease.
- Poor performances of the preconditioners $\mathbf{P}_B$ and $\mathbf{P}_T$.
- Computational costs: more inner iterations, valuation of q_{st} per inner iteration.

A Burgers system

- **Discretization** (first-order upwind scheme and second-order centered scheme):

$$\frac{1}{\Delta t}(u_i^{n+1} - u_i^n) + \frac{u_i^n}{2\Delta x}(u_{i+1}^n - u_{i-1}^n) - \frac{\nu}{(\Delta x)^2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n) = g(i\Delta x, n\Delta t)$$

with $\Delta x = 0.01$ $(n = 100)$, $\Delta t = 1.10^{-5}$, $\nu = 0.25$, $T = 0.03$ and $N_{sw} = 50$.

- **True solution**: $\begin{array}{l} x_0^t = 0.1\sin(2\pi x), \quad \forall x \in [0,1] \\ \forall j = 1 : N_{sw} \quad \mathbf{x}_j^t = \mathcal{M}_j(x_{j-1}^t) + \epsilon_j^m, \quad \epsilon_j^m \sim \mathcal{N}(0, \sigma_m^2 \mathbf{I}_n) \end{array}$ with $\sigma_m^2 = 1.10^{-4}\dfrac{T}{N_{sw}}$.

- **Observations**: $\forall j = 1 : N_{sw} \quad \mathbf{y}_j = \mathcal{H}_j(x_j^t) + \epsilon_j^o, \quad \epsilon_j^o \sim \mathcal{N}(0, 1.10^{-3}\mathbf{I}_{m_j})$, with $m_j = 20$. $\mathbf{R}_j$ diagonal such that $\kappa(\mathbf{R}_j) = 10^3$.

- **Background**: $\begin{array}{l} x_b = x_0^t + \epsilon_j^m, \quad \epsilon^b \sim \mathcal{N}(0, 1.10^{-2}\mathbf{I}_n) \\ \forall j = 1 : N_{sw} \quad \mathbf{x}_j = \mathcal{M}_j(x_{j-1}) + \epsilon_j^m, \quad \epsilon_j^m \sim \mathcal{N}(0, \sigma_m^2 \mathbf{I}_n) \end{array}$

  $\mathbf{B} = \sigma_b^2(\alpha \mathbf{I}_n + (1-\alpha)\tilde{\mathbf{B}})$, with $\tilde{B}_{i,j} = e^{-\frac{d(i,j)^2}{L^2}}$, such that $\kappa(\mathbf{B}) = 10^5$.

- **Model error**: $\mathbf{Q}_j$ similar to $\mathbf{B}$ such that $\kappa(\mathbf{Q}_j) \sim 1.610^3$.

The linearized subproblems

State Formulation

$$\min_{\delta x} \frac{1}{2}\|\mathbf{L}\delta\mathbf{x} - \mathbf{b}\|^2_{\mathbf{D}^{-1}} + \frac{1}{2}\|\mathbf{H}\delta\mathbf{x} - \mathbf{d}\|^2_{\mathbf{R}^{-1}}$$

- Matrix-vector products with $\mathbf{L}$ and $\mathbf{H}$ naturally parallel.
- Preconditioning is difficult:

$$\mathbf{P} = \mathbf{D}^{1/2}\widetilde{\mathbf{L}}^{-\mathbf{T}}(\mathbf{L}^{\mathbf{T}}\mathbf{D}^{-1}\mathbf{L})\widetilde{\mathbf{L}}^{-1}\mathbf{D}^{1/2}$$

  can be ill-conditionned depending on the accuracy of $\widetilde{\mathbf{L}}^{-1}$.

Forcing Formulation

$$\min_{\delta p} \frac{1}{2}\|\delta\mathbf{p} - \mathbf{b}\|^2_{\mathbf{D}^{-1}} + \frac{1}{2}\|\mathbf{H}\mathbf{L}^{-1}\delta\mathbf{p} - \mathbf{d}\|^2_{\mathbf{R}^{-1}}$$

- Change of variables: $\delta\mathbf{p} = \mathbf{L}\delta\mathbf{x}$
- Matrix-vector products with $\mathbf{L}^{-1}$ is a priori sequential.
- Preconditioning is straightforward: structure is similar to the strong-constrained case.

Inverse of covariance matrices: expensive operation for new systems (hybrid background error covariance matrices).

Safeguarded Saddle

Convergence

Remember $q_{st}(0) - q_{st}(\delta x) \geq \max\left(\varepsilon_q \min\left(1, \|g_k\|^2\right), \theta_j\right)$

- From the termination criterion one gets
  $\varepsilon_q \kappa_g^{-2} \|g_k\|^2 \leq -g_k^T \delta x_k - \frac{1}{2} \delta x_k^T \nabla^2 q_{\mathrm{st}}(x_k) \delta x_k$
- From the positive definiteness of $\nabla^2 q_{\mathrm{st}}$, we deduce $-g_k^T \delta x_k \geq \varepsilon_q \kappa_g^{-2} \|g_k\|^2$
- The strict convexity of q_{st} ensures that $\|\delta x_k\| \leq \frac{2}{\nu_{\min}} \|g_k\|$
- We therefore get that $-g_k^T \delta x_k \geq \kappa_1 \|g_k\|^2$ and $\|\delta x_k\| \leq \kappa_2 \|g_k\|$, in other words, δx_k is gradient related
- A cosine condition and the convergence of the linesearch naturally follows.

Some preconditioners for the saddle algorithms

Preconditioners

$$P_M = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \tilde{\mathbf{L}} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \\ \tilde{\mathbf{L}}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad P_B = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} \end{pmatrix}, \quad P_T = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \tilde{\mathbf{L}} \\ \mathbf{0} & \mathbf{R} & \mathbf{H} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} \end{pmatrix}$$

with $\mathbf{S}^{-1} = \tilde{\mathbf{L}}^{-1}\mathbf{D}\tilde{\mathbf{L}}^{-\mathsf{T}}$ and $\tilde{\mathbf{L}} \sim \mathbf{L}$ (square, nonsingular),

Inverse

$$P_M^{-1} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \tilde{\mathbf{L}}^{-1} \\ \mathbf{0} & \mathbf{R}^{-1} & \mathbf{0} \\ \tilde{\mathbf{L}}^{-\mathsf{T}} & \mathbf{0} & -\mathbf{S}^{-1} \end{pmatrix}, \quad P_B^{-1} = \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}^{-1} \end{pmatrix},$$

$$P_T^{-1} = \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} & -\mathbf{D}^{-1}\tilde{\mathbf{L}}\mathbf{S}^{-1} \\ \mathbf{0} & \mathbf{R}^{-1} & -\mathbf{R}^{-1}\mathbf{H}\mathbf{S}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}^{-1} \end{pmatrix}$$