The Iterative Solution of Systems from PDE Constrained Optimization: An Overview

Tyrone Rees

STFC Rutherford Appleton Laboratory

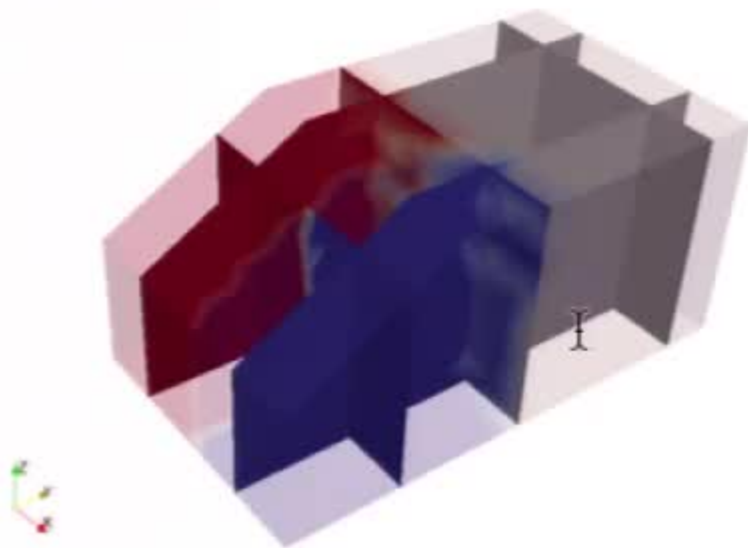
joint work with Andy Wathen (Oxford), Sue Thorne (RAL),

Martin Stoll (MPI Madgeburg), Andrew Barker (LLNL)

Alright to turn the central heating down a notch dear?

Science & Technology Facilities Council

How does this work?



$$\min_{\mathbf{y},\mathbf{u}} \frac{1}{2}\|\mathbf{y}-\hat{\mathbf{y}}\|_*^2 + \frac{\beta}{2}\|\mathbf{u}\|_*^2$$

$$\text{s.t. } \mathcal{L}\mathbf{y} = \mathbf{f}(\mathbf{u})$$

$$\mathbf{u}_l \leq \mathbf{u} \leq \mathbf{u}_u$$

$$\mathbf{y}_l \leq \mathbf{y} \leq \mathbf{y}_u$$

$\mathbf{y}$: temperature (state)

$\mathbf{u}$: control

Science & Technology Facilities Council

The ‘mother problem’

$$\min_{y,u} \frac{1}{2}||y - \hat{y}||^2_{L_2(\Omega_1)} + \frac{\beta}{2}||u||^2_{L_2(\Omega_2)}$$

$$\text{s.t.} \quad -\nabla^2 y = \begin{cases} u, & \text{for } x \in \Omega_2 \\ 0, & \text{for } x \in \Omega \backslash \Omega_2, \end{cases}$$

$$y = f \text{ on } \partial\Omega$$

$$\min_{\mathbf{y},\mathbf{u}} \frac{1}{2}\mathbf{y}^T Q_y \mathbf{y} - \mathbf{y}^T \mathbf{b} + \frac{\beta}{2}\mathbf{u}^T Q_u \mathbf{u}$$

$$\text{s.t.} \quad K\mathbf{y} = \widehat{Q}\mathbf{u} + \mathbf{f}$$

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \\ \mathbf{f} \end{bmatrix}$$

Science & Technology Facilities Council

The linear system

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \\ \mathbf{f} \end{bmatrix}$$

The linear system is:

- symmetric, but indefinite
- very large scale
- sparse

Solve with a Krylov subspace method.

Science & Technology Facilities Council

Preconditioning

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix}$$ saddle point system

Ideal preconditioner:

$$P = \begin{bmatrix} \beta Q_u & 0 & 0 \\ 0 & Q_y & 0 \\ 0 & 0 & \frac{1}{\beta}\widehat{Q} Q_u^{-1} \widehat{Q}^T + K^T Q_y^{-1} K \end{bmatrix}$$

Three distinct eigenvalues – MINRES will converge in three iterations.

[Murphy, Golub, Wathen, 1999]

Science & Technology Facilities Council

Preconditioning

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix} \quad \text{saddle point system}$$

Ideal preconditioner:

$$P = \begin{bmatrix} \beta \widehat{Q_u} & 0 & 0 \\ 0 & \widehat{Q_y} & 0 \\ 0 & 0 & \widehat{\frac{1}{\beta}\widehat{Q}Q_u^{-1}\widehat{Q}^T + K^T Q_y^{-1} K} \end{bmatrix}$$

Three distinct eigenvalues – MINRES will converge in three iterations.

[Murphy, Golub, Wathen, 1999]

Science & Technology Facilities Council

Approximating $\begin{bmatrix} \beta Q_u & 0 \\ 0 & Q_y \end{bmatrix}$

The diagonal is a good approximation to the mass matrix: e.g., if we have a 3D tetrahedral mesh of P1 elements we have

$$\lambda(D^{-1}Q) \in [1/2, 5/2].$$

[Wathen, 1987]

Can do better – Chebyshev semi-iteration applied to relaxed Jacobi.

[Wathen, R., 2008] , [R., Dollar, Wathen, 2008]

Science & Technology Facilities Council

Approximating $\frac{1}{\beta}\widehat{Q}Q_u^{-1}\widehat{Q}^T + K^T Q_y^{-1} K$

Approximation 1: $S \approx K^T Q_y^{-1} K$

Eigenvalues of the preconditioned system satisfy:

$$\begin{aligned} \lambda &= 1, \\ \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right) \leq \lambda &\leq \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \\ \text{or } \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \leq \lambda &\leq \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right). \end{aligned}$$

(h mesh size, α_1, α_2 constants)

[R., Dollar, Wathen, 2008]

Science & Technology Facilities Council

Approximating $\frac{1}{\beta}\widehat{Q}Q_u^{-1}\widehat{Q}^T + K^T Q_y^{-1} K$

Approximation 2: $S \approx (K^T + L^T)Q_y^{-1}(K + L)$, where
$L^T = \frac{1}{\sqrt{\beta}}\widehat{Q}Q_u^{-1/2}Q_y^{1/2}$

Eigenvalues of the preconditioned system satisfy:

$$\frac{1}{2} \leq \lambda \leq 1$$

[Pearson, Wathen, 2011]

Control constraints

To make the problem a little harder...

$$\begin{aligned} \min_{y,u} \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega_1)} &+ \frac{\beta}{2}\|u\|^2_{L_2(\Omega_2)} \\ \text{s.t.} \qquad -\nabla^2 y &= u \text{ in } \Omega_2 \\ y &= f \text{ on } \partial\Omega \\ u_l \leq u &\leq u_u \end{aligned}$$

Can be solved using an active set method (semi-smooth Newton).

Linear system at each step looks like:

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mu_k \\ \mathbf{b} \\ \mathbf{f} \end{bmatrix}$$

[Stoll, Wathen, 2012]

State constraints

$$\min_{y,u} \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega_1)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega_2)} + \frac{1}{2\epsilon}\|\max\{0, y-y_u\}\|^2_{L_2(\Omega)} + \frac{1}{2\epsilon}\|\min\{0, y-y_l\}\|^2_{L_2(\Omega)}$$

$$\text{s.t.} \quad -\nabla^2 y = u \text{ in } \Omega_2$$

$$y = f \text{ on } \partial\Omega_2$$

Need to solve a linear system of the form

$$\begin{bmatrix} \beta Q_u & 0 & -\widehat{Q}^T \\ 0 & Q_y + \epsilon^{-1} G_{\mathcal{A}} Q_y G_{\mathcal{A}} & K^T \\ -\widehat{Q} & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} - \mathbf{z} \\ \mathbf{f} \end{bmatrix}$$

where G is a projection onto the active set $\mathcal{A}$

[Ito, Kunisch, 2003] , [Pearson, Stoll, Wathen, 2014]

Science & Technology Facilities Council

Time-dependent problems

$$\min_{y,u} \frac{1}{2}\int_0^T \int_{\Omega_1} (y-\hat{y})^2 dxdt + \frac{\beta}{2}\int_0^T \int_{\Omega_2} u^2 dxdt$$

$$\text{s.t.} \qquad y_t - \triangle y = \begin{cases} u, & \text{for } (\mathbf{x}, t) \in \Omega_2 \times [0, T], \\ 0, & \text{for } (\mathbf{x}, t) \in \Omega \backslash \Omega_2 \times [0, T], \end{cases}$$

$$y = g, \quad \text{on } \partial\Omega,$$

$$y = y_0, \quad \text{at } t = 0,$$

Science & Technology Facilities Council

Transformed into linear algebra...

$$\begin{bmatrix} \tau\beta\mathcal{M}_u & 0 & -\tau\mathcal{M} \\ 0 & \tau\mathcal{M}_y & \mathcal{K}^T \\ -\tau\mathcal{M} & \mathcal{K} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 \\ \tau\widehat{\mathbf{b}} \\ \mathbf{d} \end{bmatrix},$$

$\mathcal{M} = \mathrm{blkdiag}(\widehat{Q}, \ldots, \widehat{Q})$,

$\mathcal{M}_y = \mathrm{blkdiag}(1/2Q_y, Q_y, \ldots, Q_y, 1/2Q_y)$,

$\mathcal{M}_u = \mathrm{blkdiag}(1/2Q_u, Q_u, \ldots, Q_u, 1/2Q_u)$, and

$$\mathcal{K} = \begin{bmatrix} Q_u + \tau K & & & \\ -Q_u & Q_u + \tau K & & \\ & \ddots & \ddots & \\ & & -Q_u & Q_u + \tau K \end{bmatrix}.$$

[Pearson, Stoll, Wathen, 2012]

Science & Technology Facilities Council

Other norms?

$$\min_{y,u} \frac{1}{2}\|y-\widehat{y}\|^2_{L_2(\Omega_1)} + \frac{\beta}{2}\|u\|^2_{\mathcal{H}_1(\Omega_2)}$$
$$\text{s.t.} \quad -\nabla^2 y = u \text{ in } \Omega_2$$
$$y = f \text{ on } \partial\Omega$$

Science & Technology Facilities Council

Time-dependent formulation

$$\min_{y,u} \frac{1}{2}\int_0^T \int_{\Omega_1} (y-\bar{y})^2 dxdt + \frac{\beta}{2}\int_0^T \int_{\Omega_2} u^2 dxdt + \frac{\beta}{2}\int_0^T \int_{\Omega_2} (\nabla u)^2 dxdt$$

$$y_t - \nabla^2 y = \begin{cases} u, & \text{for } (\mathbf{x}, t) \in \Omega_2 \times [0, T], \\ 0, & \text{for } (\mathbf{x}, t) \in \Omega \backslash \Omega_2 \times [0, T], \end{cases}$$

$$y = g, \quad \text{on } \partial\Omega,$$

$$y = y_0, \quad \text{at } t = 0,$$

Science & Technology Facilities Council

The linear system

$$\begin{bmatrix} \tau\mathcal{M}_y & 0 & -\mathcal{K}^T \\ 0 & \tau\beta(\mathcal{M}_u + \mathcal{K}_u) & \tau\widehat{\mathcal{M}} \\ -\mathcal{K} & \tau\widehat{\mathcal{M}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \tau\widehat{\mathbf{b}} \\ 0 \\ \mathbf{d} \end{bmatrix}$$

where $\mathcal{K}_u = \text{blkdiag}(1/2K_u, K_u, \ldots, K_u, 1/2K_u)$, K_u is a Neumann Laplacian.

Approximation to $S = \tau^{-1}\mathcal{K}\mathcal{M}_y^{-1}\mathcal{K} + \tau\beta^{-1}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\widehat{\mathcal{M}}^T$.

Look for a (non-symmetric) approximation

$$\hat{S} = \tau^{-1}(\mathcal{K} + \mathcal{L}_1)\mathcal{M}_y^{-1}(\mathcal{K} + \mathcal{L}_2)^T$$

Approximation to $S = \tau^{-1}\mathcal{K}\mathcal{M}_y^{-1}\mathcal{K} + \tau\beta^{-1}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\widehat{\mathcal{M}}^T$.

Look for a (non-symmetric) approximation

$$\hat{S} = \tau^{-1}\left(\mathcal{K} + \frac{\tau}{\sqrt{\beta}}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\mathcal{M}_y\right)\mathcal{M}_y^{-1}\left(\mathcal{K} + \frac{\tau}{\sqrt{\beta}}\widehat{\mathcal{M}}\right)^T$$

Science & Technology Facilities Council

Approximation to $S = \tau^{-1}\mathcal{K}\mathcal{M}_y^{-1}\mathcal{K} + \tau\beta^{-1}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\widehat{\mathcal{M}}^T$.

Look for a (non-symmetric) approximation

$$\hat{S} = \tau^{-1}\left(\mathcal{K} + \frac{\tau}{\sqrt{\beta}}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\mathcal{M}_y\right)\mathcal{M}_y^{-1}\left(\mathcal{K} + \frac{\tau}{\sqrt{\beta}}\widehat{\mathcal{M}}\right)^T$$

Solving a system with $\mathcal{K} + \frac{\tau}{\sqrt{\beta}}\widehat{\mathcal{M}}(\mathcal{M}_u + \mathcal{K}_u)^{-1}\mathcal{M}_y\mathbf{u} = \mathbf{f}$ is equivalent to solving

$$\begin{bmatrix} \mathcal{K} & \widehat{\mathcal{M}} \\ \mathcal{M}_y & -\frac{\sqrt{\beta}}{\tau}(\mathcal{M}_u + \mathcal{K}_u) \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ * \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}.$$

Treat this as a sub-problem.

Science & Technology Facilities Council

Approximating a solve with $\begin{bmatrix} \mathcal{K} & \widehat{\mathcal{M}} \\ \mathcal{M}_y & -\frac{\sqrt{\beta}}{\tau}(\mathcal{M}_u + \mathcal{K}_u) \end{bmatrix}$

Recall this involves the solution of diagonal blocks of the form

$$\begin{bmatrix} Q_u + \tau K & \widehat{Q} \\ Q_y & -\frac{\sqrt{\beta}}{\tau}(Q_u + K_u) \end{bmatrix}$$

Approximate this by a simple iteration of the form

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \omega W^{-1}\mathbf{r}_k$$

where $W = \begin{bmatrix} \widehat{M + \tau K} & 0 \\ 0 & -\frac{\sqrt{\beta}}{\tau}(\widehat{M_u + K_u}) \end{bmatrix}$

Science & Technology Facilities Council

Numerical results

Example:

$$\hat{y} = \exp\left(-64\left((x_0 - 0.5)^2 + (x_1 - 0.5)^2\right)\right),$$

$\Omega = [0,1]^2$, $\tau = 0.05$. In the preconditioner $\omega = 0.1$, 10 steps of simple iteration taken. Outer tolerance is 10^{-6}.

DoF	$\beta = 10^{-2}$ # it(t)	$\beta = 10^{-4}$ # it(t)	$\beta = 10^{-6}$ # it(t)
1089	13(35.1)	13(35.2)	22(57.3)
4225	13(112.6)	15(128.8)	22(184.8)
16641	15(462.3)	15(462.2)	25(756.1)
66049	17(1442.6)	20(1691.4)	31(2578.7)
263169	19(4928.3)	22(5843.9)	34(8368.3)

[Barker, R., Stoll, to appear]

Science & Technology Facilities Council