Efficient preconditioning of optimality systems with non-self-adjoint state operators

Kent-Andre Mardal and Bjørn Fredrik Nielsen

Center for Biomedical Computing

10th September 2013

The PDE constrained optimization problem

Consider the following optimization problem:

$$\min_{v,u \in V,U} \|u - d\|_K^2 + \alpha \|v - v_{prior}\|_L^2$$

subject to

$$Au = Bv.$$

Here, A is a PDE that is not self-adjoint!

And,

$$\|\cdot\|_K^2 = K(\cdot,\cdot) \text{ and } \|\cdot\|_L^2 = L(\cdot,\cdot),$$

where $K : U \times U \to \mathbb{R}$ is a symmetric and positive semi-definite bilinear form, while $L : V \times V \to \mathbb{R}$ is a symmetric and positive definite bilinear form.

The regularization parameter α is usually small causes an ill-conditioned/ill-posed problem

R_X is the Riesz mapping between X and X'.

Preconditioning Approach for non-self-adjoint operators

Using the method of Larange multipliers we arrive at the following problem:

$$A_\alpha \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} v_{prior} \\ Kd \\ 0 \end{bmatrix}$$

Here, $u \in U$, $v \in V$, and $w \in W$.

We propose the following preconditioners:

$$B_1 = \begin{bmatrix} \alpha L & 0 & 0 \\ 0 & K + \alpha A' R_W A & 0 \\ 0 & 0 & \frac{1}{\alpha} B L^{-1} B' \end{bmatrix},$$

$$B_2 = \begin{bmatrix} \alpha L & 0 & 0 \\ 0 & K & 0 \\ 0 & 0 & \frac{1}{\alpha} B L^{-1} B' + A K^{-1} A' \end{bmatrix},$$

$$B_3 = \begin{bmatrix} R_{V'} & 0 & 0 \\ 0 & R_{U'} & 0 \\ 0 & 0 & R_{W'} \end{bmatrix}.$$

- B_1 : is based on Nielsen and Mardal, SIAM J. Control Optim., 2010 and related to Schöberl and Zulehner, Numer. Math 2007
- B_2 : is a preconditioner based on the Schur complement
- B_3 : is based on Nielsen and Mardal, SISC, 2013

Assumptions B_1

We always assume that:

(A1) $A : U \to W'$ is an isomorphism,

(A2) $L : V \to V'$ is an isomorphism,

(A3) $K : U \to U'$ is bounded,

(A4) $B : V \to W'$ is bounded.

For B_1 we will need the additional assumption:

(A5) $c_0 (Lv, v) \leqslant (R_w Bv, Bv) \leqslant c_1 (Lv, v), \quad \forall v \in V.$

(A6) $(L^{-1} B' w, B' w) > 0, \quad \forall w.$

B_1: Coersivity condition

Coersivity follows since, for v and u such that $Bv + Au = 0 \in W'$ we have that

$$(R_W Au, Au) = (R_W Bv, Bv).$$

Therefore,

$$\begin{aligned}
(\alpha Lv, v) + (Ku, u) &= \frac{1}{2}(\alpha Lv, v) + \frac{1}{2}(\alpha Lv, v) + (Ku, u) \\
&\geqslant \frac{1}{2c_1}(\alpha R_W Bv, Bv) + \frac{1}{2}(\alpha Lv, v) + (Ku, u) \\
&\geqslant \frac{1}{2c_1}(\alpha R_W Au, Au) + \frac{1}{2}(\alpha Lv, v) + (Ku, u)
\end{aligned}$$

Since we assumed that:

$$c_0(Lv, v) \leqslant (R_w Bv, Bv) \leqslant c_1(Lv, v)\ \forall v \in V$$

B_1: The inf-sup condition

The inf-sup condition also follows.
Given w, let $\hat{u} = 0$ and $\hat{v} = L^{-1}B'w$:

$$\begin{aligned}
\sup_{v,u} & \frac{(Bv + Au, w)}{((\alpha Lv, v) + ((K + \alpha A' R_W A)u, u))^{1/2}} \\
\geqslant & \frac{(B\hat{v} + A\hat{u}, w)}{((\alpha L\hat{v}, \hat{v}) + ((K + \alpha A' R_W A)\hat{u}, \hat{u}))^{1/2}} \\
& \frac{(L^{-1}B'w, B'w)}{(\alpha L L^{-1}B'w, L^{-1}B'w)^{1/2}} \\
& = (\alpha^{-1}L^{-1}B'w, B'w)^{1/2}
\end{aligned}$$

Hence, the inf-sup condition is valid.

Assumptions B_2

We will always assume that:

(A1) $A : U \to W'$ is an isomorphism,

(A2) $L : V \to V'$ is an isomorphism,

(A3) $K : U \to U'$ is bounded,

(A4) $B : V \to W'$ is bounded.

And for B_2 we need that U, V, W are contained in a larger space X such that

- $V \subset X \subset X' \subset V'$
- $U \subset X \subset X' \subset U'$
- $W \subset X \subset X' \subset W'$

And that:

(A7) $K : X \to X'$ is an isomorphism,

(A8) $A : X \to W'$ is an bounded,

(A9) $\frac{1}{\alpha} BL^{-1}B' + AK^{-1}A' > 0$.

The B_2 preconditioner

We remember the operator:

$$A_\alpha \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} v_{prior} \\ Kd \\ 0 \end{bmatrix}$$

and the preconditioner B_2:

$$B_2 = \begin{bmatrix} \alpha L & 0 & 0 \\ 0 & K & 0 \\ 0 & 0 & \frac{1}{\alpha} BL^{-1}B' + AK^{-1}A' \end{bmatrix},$$

The B_2 preconditioner is well-defined since $\ker(L)$ and $\ker(K)$ are zero. and the Schur complement $\frac{1}{\alpha} BL^{-1}B' + AK^{-1}A'$ is well-defined. This preconditioner is therefore effective, c.f. Murphy, Golub, Wathen, SISC 2000.

The preconditioner B_3

We remember the system matrix:

$$\begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix}$$

and the preconditioner

$$B_3 = \begin{bmatrix} R_{V'} & 0 & 0 \\ 0 & R_{U'} & 0 \\ 0 & 0 & R_{W'} \end{bmatrix}.$$

As described in Nielsen and Mardal, SISC, 2013, it is sufficient for an efficient MinRes algorithm that the eigenvalues are in bounded intervals, i.e. $(-C, -c)$, $(\alpha, 2\alpha)$, (c, C) + some outliers

The proof extends nicely to the situation where A is non-self-adjoint.

ODE Example

Find u, v such that

$$\int_\Omega k_o(u-d)^2 \rightarrow \min$$

such that $u_t + au = v$.

$$a = At^2, \quad k_o = 1$$

Assume $u \in H_0^1$, $v, w \in L^2$ and solve

$$A_\alpha \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha I & 0 & I' \\ 0 & I & (\frac{\partial}{\partial t})' \\ I & \frac{\partial}{\partial t} & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} v_{prior} \\ d \\ 0 \end{bmatrix}$$

Preconditioners for the ODE example

$$B_1 = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & I + \alpha(-\Delta - a' + a^2) & 0 \\ 0 & 0 & \frac{1}{\alpha} I \end{bmatrix}$$

$$B_2 = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \frac{1}{\alpha} I + (-\Delta_h + a' + a^2) \end{bmatrix},$$

$$B_3 = \begin{bmatrix} I & 0 & 0 \\ 0 & -\Delta & 0 \\ 0 & 0 & I \end{bmatrix}.$$

Not that the Lagrange multiplier consists of piecewise constant elements, therefore we use the following $-\Delta$ operator:

$$(-\Delta_h w_h, l_h) = \sum_{i=0}^{N-1} [w_h]_{x_i} [l_h]_{x_i}, \quad \forall w_h, l_h \in W_h.$$

Condition numbers for ODE Example

	B_1A			B_2A			B_3A		
$N\backslash\alpha$	1.0	1.0e-3	1.0e-6	1.0	1.0e-3	1.0e-6	1.0	1.0e-3	1.0e-6
20	6.0	6.0	5.8	13.8	5.9	5.8	6.0	1650	9382
100	6.0	6.0	5.8	14.3	6.3	5.8	6.0	1983	2.1e+5
500	6.0	6.0	6.0	14.4	6.5	5.9	6.0	2000	1.5e+6

Table: Condition number using CG and DG elements.

The condition numbers seem nicely bounded for B_1 and B_2. For B_3 the eigenvalues are nicely bounded except for an interval of eigenvalues around α.

MinRes Iterations for ODE Example

	B_1A			B_2A			B_3A		
$N \backslash \alpha$	1.0	1.0e-3	1.0e-6	1.0	1.0e-3	1.0e-6	1.0	1.0e-3	1.0e-6
20	27	45	8	42	27	8	27	67	96
100	33	55	14	62	33	12	33	75	324
500	41	52	38	77	32	22	41	63	504
2500	58	46	48	52	28	28	58	47	266
12500	36	42	46	59	26	26	36	42	108

Table: Number of iterations using CG and DG elements. Convergence criteria is that the preconditioned residual is reduced by $1.0e - 10$.

Second ODE example

We consider the case with only one observation, i.e.,

$$k_o^{\delta}(t) = \begin{cases} 1 & \text{for } t = 0.5, \\ 0 & \text{elsewhere.} \end{cases} \tag{1}$$

Furhermore, we let

$$V = \text{span}\{t^i\} \quad i = 0, M-1.$$

Second ODE example: MinRes iterations

	$B_1 A$			$B_3 A$		
$N \backslash \alpha$	1.0	1.0e-3	1.0e-6	1.0	1.0e-3	1.0e-6
20	26	35	23	26	51	68
100	33	44	38	33	63	188
500	42	49	62	42	62	240
2500	58	62	76	58	56	117
12500	35	43	60	36	34	48

Table: Number of iterations using CG and DG elements with observations only in parts of the domain as defined in (**??**) and a lower order *a* defined as above. Convergence criteria is that the preconditioned residual is reduced by $1.0e - 10$.

Perverted Poisson

Find u, v such that

$$\int_\Omega (u-d)^2 \to \min$$

such that

$$-\Delta u = v.$$

Assume $u \in H_0^2$ and $v, w \in L^2$ and solve

$$A_\alpha \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha I & 0 & I' \\ 0 & I & -\Delta' \\ I & \Delta & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} v_{prior} \\ d \\ 0 \end{bmatrix}$$

A suitable preconditioner is:

$$B = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & I - \alpha\Delta^2 & 0 \\ 0 & 0 & \frac{1}{\alpha} I \end{bmatrix}$$

u discretized by Hermite elements and v, w by first order DG elements

Condition number for the 'Perverted Poisson porblem'

	$B_1 A$		
$N \backslash \alpha$	1.0	1.0e-2	1.0e-4
8	4.0	4.0	3.9
16	4.0	4.0	4.0
32	4.0	4.0	4.0
64	4.0	4.0	4.0
128	4.0	4.0	4.0

Table: Condition number for the Poisson (in $L(H^2, L^2)$) problem using Hermite elements combined with discontinuous Lagrange for the Lagrange multipiler.

[Thanks to Adrian Hope for conducting these experiments]

Transport state equation

Let us first consider the following problem:
Find $u \in V$ and $v \in L^2$ such that

$$\int_0^T (u - u_d)^2 + \alpha(v - v_{prior})^2 \, dt \rightarrow \min$$

and

$$u_t - u_x = v, \quad t \in [0, T].$$

Note that for $Au = u_t - u_x$ we have that

$$A'A = -\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial t \partial x} + \frac{\partial^2}{\partial x \partial t} - \frac{\partial^2}{\partial x^2}$$

Hence, $A'A$ is not spectrally equivalent with $-\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}$ and $A'A$ has a large kernel consisting of functions u such that $u_t = u_x$. Let us therefore introduce the space V with an inner product defined by

$$(u, v)_V = (u, v) + (u_t, v_t) - (u_x, v_t) - (u_t, v_x) + (u_x, u_v).$$

We assume Dirichlet boundary conditions on the whole boundary.

Preconditioner

Assume $u \in V_0$ and $v, w \in L^2$ and solve

$$A_\alpha \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha I & 0 & I' \\ 0 & I & A' \\ I & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} v_{prior} \\ d \\ 0 \end{bmatrix}$$

A suitable preconditioner is:

$$B = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & I - \alpha A A' & 0 \\ 0 & 0 & \frac{1}{\alpha} I \end{bmatrix}$$

Using Exact Preconditioning

	B_1A			B_3A		
$N \backslash \alpha$	1.0	1.0e-2	1.0e-4	1.0	1.0e-2	1.0e-4
4	4.0	3.6	2.6	4.0	242	1650
8	4.0	4.0	2.6	4.0	273	8010
16	4.0	4.0	2.7	4.0	281	30650

Table: Condition number for transport state equation.

Preconditioner using ILU

	B_1A			B_3A		
$N\setminus\alpha$	1.0	1.0e-2	1.0e-4	1.0	1.0e-2	1.0e-4
32	38	49	23	124	65	74
64	50	61	32	220	103	65
128	68	84	40	390	183	48
256	92	121	43	784	322	32

Table: Number of iterations for the transport equation using CG and DG elements with observations in the whole the domain. Approximate preconditioner constructed using ILU. Convergence criteria is that the preconditioned residual is reduced by $1.0e - 10$.

Preconditioner using ILU

	B_1A			B_3A		
$N\backslash\alpha$	1.0	1.0e-2	1.0e-4	1.0	1.0e-2	1.0e-4
32	60	44	23	97	65	97
64	78	57	32	127	89	133
128	104	75	39	167	115	163
256	128	99	41	199	145	191

Table: Number of iterations for the transport equation using CG and DG elements with observations in the whole the domain. Approximate preconditioner constructed using ML. Convergence criteria is that the preconditioned residual is reduced by $1.0e-10$.

Conclusion

- We have successfully developed efficient preconditioners for a series of PDE constrained optimization problems with non-self-adjoint state operators
- Our technique often requires additional boundary conditions on the state (which may be an advantage) and extra regularity
- Non-standard Hilbert spaces are imporant
- This technique gives well-posed problems (in non-standard spaces) even without observations everywhere
- The work is still work in progress and has not been published yet