Preconditioners for Elliptic Control Problems

Ekkehard W. Sachs

Interdisciplinary Center for Applied Mathematics (ICAM)
Department of Mathematics, Virginia Tech
and
Department of Mathematics, Universität Trier, Germany

KKT Seminar

Special Radon Semester, RICAM, Linz

November 30, 2005

Preconditioning KKT Systems

General Formulation of Problem

$$f : R^n \to R, \qquad h : R^n \to R^m$$

f,h twice differentiable

$$\text{Minimize } f(x) \quad \text{with} \quad h(x) = 0.$$

Optimal control problem

$$x = (y, u), \qquad y \in Y, \; u \in U$$

System equation (y state, u control, design)

$$h(x) = h(y, u) = y - S(y, u) = 0$$

Lagrangian

$$L(x; l) = f(x) + l^T h(x), \quad x \in R^n, \; l \in R^m,$$

Necessary Optimality Conditions

Theorem

If a constraint qualification holds at optimum x_*, then there exists a Lagrange multiplier λ_* such that

$$\begin{aligned} L_x(x_*, \lambda_*) &= \nabla f(x_*) - J(x_*)^T \lambda_* &= 0 \\ L_\lambda(x_*, \lambda_*) &= h(x_*) &= 0 \end{aligned}$$

Sequential **Q**uadratic **P**rogramming similar to Newton's method.
Need to solve **KKT systems**, linear systems of type

$$\begin{pmatrix} L_{xx}(x,\lambda) & J(x) \\ J(x)^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} L_x(x,\lambda) \\ h(x) \end{pmatrix}$$

Preconditioning KKT Systems

Splitting of variables: $x = (y, u)$, (state, control or design)

Construction of Preconditioners for linear systems with system matrix

$$K = \begin{pmatrix} L_{yy} & L_{yu} & A^T \\ L_{uy} & L_{uu} & B^T \\ A & B & 0 \end{pmatrix},$$

where

$$\begin{array}{ll} L_{yy} \in R^{m \times m}, & L_{uu} \in R^{k \times k}, \\ L_{yu} \in R^{m \times k}, & L_{uy} \in R^{k \times m}, \\ A \in R^{m \times m}, & B \in R^{m \times k}. \quad (k \ll m) \end{array}$$

An Indefinite Preconditioner for KKT systems Arising in Optimal Control Problems

A. Battermann, EWS   2002

Elliptic Boundary Value Problem

Test Problem

$$\min_{(y,u)} \Big\{ \beta_1 \int_{\Gamma_2} Q_2^2(y(x,\bar{z}),u(x))\,dx + \beta_2 \int_{\Gamma_2} (u(x)-u_v)^2\,dx + \beta_3 \int_{\Gamma_6} Q_6(y(0,z))\,(y_f - y_6)\,dz \Big\}$$

s.t.

$$\begin{aligned}
\Delta y(x,z) &= 0 && \text{in } \Omega, \\
\tfrac{\partial}{\partial n} y(x,z) &= 0 && \text{on } \Gamma_1 \cup \Gamma_3 \cup \Gamma_5, \\
y(\bar{x},z) &= y_4 && \text{on } \Gamma_4, \\
y(0,z) &= y_6 && \text{on } \Gamma_6, \\
\tfrac{\partial}{\partial n} y(x,\bar{z}) &= \tfrac{1}{d}\,(u(x) - y(x,\bar{z})) && \text{on } \Gamma_2.
\end{aligned}$$

Industrial partner: TGU, Koblenz. (Groundwater modelling)

Elliptic Control Problem

Battermann, Heinkenschloss (1996)

$$\text{Min} \qquad \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\gamma}{2}\int_{\partial\Omega} u^2(s) ds$$

subject to

$$\begin{aligned} -\Delta y(x) + y(x) &= f(x) \quad & x \in \Omega \\ \frac{\partial}{\partial n} y(x) &= u(x) \quad & x \in \partial\Omega \end{aligned}$$

and

$$\begin{aligned} y_{low} \le y(x) \le y_{upp} \quad & \text{a.e. in } \Omega \\ u_{low} \le u(x) \le u_{upp} \quad & \text{a.e. in } \partial\Omega \end{aligned}$$

Structure of KKT Matrix

$$K = \begin{pmatrix} L_{yy} & L_{yu} & A^T \\ L_{uy} & L_{uu} & B^T \\ A & B & 0 \end{pmatrix} = \begin{pmatrix} M_y & 0 & A^T \\ 0 & M_u & B^T \\ A & B & 0 \end{pmatrix}$$

where

M_y	is mass matrix for states
M_u	is mass matrix for controls
A	is stiffness matrix from PDE
B	is (boundary) control input matrix

Ill–Conditioning

Iterations of GMRES on original system K.
(In all computations, $n_x = 2n_z$ and $k = n$.)

grid size	4	8	16	32
dimension	94	314	1138	4322
#it	8	295	1031	3620

Condition numbers of original system K and submatrix A for different grid sizes.

grid size	4	8	16	32
dimension	94	314	1138	4322
$\kappa(K)$	1356	11238	94713	832836
$\kappa(A)$	48	192	770	3083

Left Preconditioner

$$\tilde{K} = \begin{pmatrix} 0 & 0 & \tilde{A}^T \\ 0 & \tilde{H} & B^T \\ \tilde{A} & B & 0 \end{pmatrix}$$

$\tilde{A}$ preconditioner for A, $\tilde{H}$ preconditioner for

$$H = B^T A^{-T} M_y A^{-1} B + M_u$$

Sometimes $\tilde{H} = M_u$ is sufficient.

$\tilde{K}$ block-triangular, reasonable solves for $\tilde{K}x = -r$

Cost: System and adjoint solve plus solve for H.

Results for GMRES

Iterations of GMRES on preconditioned system $\tilde{K}^{-1}K$.

grid size	4	8	16	32	64	128
dimension	94	314	1138	4322	16834	66434
$\tilde{A} = A, \tilde{H} = H$	3	3	3	3	3	3
$\tilde{A} = A, \tilde{H} = M_u$	5	4	4	4	4	4
$\tilde{A} = A, \tilde{H} = I$	4	5	5	4	4	4
$\tilde{A} = ILU(10^{-4}), \tilde{H} = H$	3	4	4	6	10	16
$\tilde{A} = ILU(10^{-3}), \tilde{H} = H$	4	5	7	11	20	37
$\tilde{A} = ILU(10^{-4}), \tilde{H} = I$	5	7	9	10	15	24
$\tilde{A} = ILU(10^{-3}), \tilde{H} = I$	5	9	13	17	28	50

Convergence Analysis

Theorem (Saad and Schultz)

For the ideal case, i.e. $\tilde{A} = A, \tilde{H} = H$, the minimal polynomial of $\tilde{K}^{-1}K$ has degree 3; hence GMRES terminates after 3 steps.

Convergence analysis in non-ideal case difficult.

Therefore consider preconditioners which preserve symmetry.

Block Preconditioners for KKT Systems in PDE-Governed Optimal Control Problems

Battermann, EWS 2003

Preconditioner 1

$$\tilde{K}_1^{-1} = \begin{pmatrix} M_y^{-1/2} & 0 & 0 \\ 0 & M_u^{-1/2} & 0 \\ 0 & 0 & M_y^{1/2}\tilde{A}^{-1} \end{pmatrix}$$

Cost: System and adjoint solve.

This yields the iteration matrix

$$\tilde{K}_1^{-1} K \tilde{K}_1^{-T} = \begin{pmatrix} I & 0 & M_y^{-1/2}A^T\tilde{A}^{-T}M_y^{1/2} \\ 0 & I & M_u^{-1/2}B^T\tilde{A}^{-T}M_y^{1/2} \\ M_y^{1/2}\tilde{A}^{-1}AM_y^{-1/2} & M_y^{1/2}\tilde{A}^{-1}BM_u^{-1/2} & 0 \end{pmatrix}$$

and in the ideal case $\tilde{A} = A$

$$\begin{pmatrix} I & 0 & I \\ 0 & I & M_u^{-1/2}B^TA^{-T}M_y^{1/2} \\ I & M_y^{1/2}A^{-1/2}BM_u^{-1} & 0 \end{pmatrix}$$

Analysis of Preconditioner 1

$$K_1 = \tilde{K}_1^{-1} K \tilde{K}_1^{-T} = \begin{pmatrix} I & 0 & I \\ 0 & I & G^T \\ I & G & 0 \end{pmatrix}$$

with $G = M_y^{1/2} A^{-1} B M_u^{-1/2}$.

Note that $I + G^T G$ has same eigenvalues as $M_u^{-1} H$.

Theorem (Battermann, EWS)

Denote the k eigenvalues of $G^T G$ by λ_j. The eigenvalues μ_i of K_1 are

$$\begin{aligned} \mu_i &= 1 & i &= 1, \dots, k \\ \mu_i &= 1/2(1 \pm \sqrt{5}) & i &= k+1, \dots, 2m-k \\ \mu_i &= 1/2(1 \pm \sqrt{5 + 4\lambda_i}) & i &= 2m-k+1, \dots, 2m+k \end{aligned}$$

Since $\lambda_i \in [O(h^p), O(1)]$, mesh independence of eigenvalues.

Preconditioner 2

$$\tilde{K}_2^{-1} = \begin{pmatrix} M_y^{-1/2} & 0 & 0 \\ 0 & M_u^{-1/2} & 0 \\ -M_y^{-1/2} & -M_y^{1/2}\tilde{A}^{-1}BM_u^{-1} & M_y^{1/2}\tilde{A}^{-1} \end{pmatrix}$$

Cost: System and adjoint solve.

This yields iteration matrix

$$\tilde{K}_2^{-1}K\tilde{K}_2^{-T} = \begin{pmatrix} I & 0 & -I + C^T \\ 0 & I & 0 \\ -I + C & 0 & I - C - C^T - GG^T \end{pmatrix}$$

with

$$\begin{aligned} G &= M_y^{1/2}\tilde{A}^{-1}BM_u^{-1/2} \\ C &= M_y^{1/2}\tilde{A}^{-1}AM_y^{-1/2} \end{aligned}$$

Analysis Preconditioner 2

In ideal Case $\tilde{A} = A$ implies $C_1 = C_2 = I$

$$K_2 = \tilde{K}_2^{-1} K \tilde{K}_2^{-T} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & -I - GG^T \end{pmatrix}$$

Theorem (Battermann, EWS)

Denote the k eigenvalues of GG^T by λ_j. The eigenvalues μ_i of K_2 are given by

$$\begin{aligned} \mu_i &= -1 - \lambda_i \quad & i &= 1, \dots, k \\ \mu_i &= -1 & i &= k+1, \dots, m \\ \mu_i &= 1 & i &= m+1, \dots, 2m+k \end{aligned}$$

since $\lambda_i \in [O(h^p), O(1)]$, mesh independence of eigenvalues.

Outlook Preconditioner 2

$$\tilde{K}_2^{-1} K \tilde{K}_2^{-T} = \begin{pmatrix} I & 0 & \Delta^T \\ 0 & I & 0 \\ \Delta & 0 & -I - GG^T - \Delta - \Delta^T \end{pmatrix}$$

with

$$\begin{aligned} G &= M_y^{1/2} A^{-1} B M_u^{-1/2} \\ \Delta &= M_y^{1/2} (\tilde{A}^{-1} A - I) M_y^{-1/2} \end{aligned}$$

Use Bramble-Pasciak preconditioner and apply CG method.
Alternative to CG with constraint preconditioner (projected CG).

Preconditioner 3

Consider ideal case $\tilde{A} = A$

$$\tilde{K}_3^{-1} = \begin{pmatrix} -B^T A^{-T} & I & B^T A^{-T} M_y A^{-1} \\ 0 & 0 & A^{-1} \\ I & 0 & -\frac{1}{2} M_y A^{-1} \end{pmatrix}$$

Cost: 2 system and **2** adjoint solves.

This yields iteration matrix

$$\tilde{K}_3^{-1} K \tilde{K}_3^{-T} = \begin{pmatrix} H & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix}$$

with $H = M_u + B^T A^{-T} M_y A^{-1} B$.

Additional preconditioning of 1-1-block yields $I + G^T G$ instead of H.

Analysis Preconditioner 3

Ideal Case $\tilde{A} = A$ and preconditioning with M_u.

$$K_3 = \begin{pmatrix} I + G^T G & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix}$$

Theorem (Battermann, EWS)

Denote the k eigenvalues of GG^T by λ_j. The eigenvalues μ_i of K_3 are given by

$$\begin{aligned} \mu_i &= 1 & i &= 1, \dots, m \\ \mu_i &= -1 & i &= m+1, \dots, 2m \\ \mu_i &= 1 + \lambda_i & i &= 2m+1, \dots, 2m+k \end{aligned}$$

Since $\lambda_i \in [O(h^p), O(1)]$, mesh independence of eigenvalues.

Numerical Comparison (Iteration Count)

GMRES for K and $\tilde{K}^{-1}K$ with $P_A = A, P_H = H$ and $P_A = A, P_H = I$.
MINRES for precond. K_1, K_2 and K_3 with $\tilde{A} = A$.

n_z	4	8	16	32	64	128
N	94	314	1138	4322	16834	66434
K	88	294	1028	3666	$\star$	$\star$
$\tilde{K}\ (P_A = A, P_H = H)$	3	3	3	3	3	3
$\tilde{K}\ (P_A = A, P_H = I)$	4	4	4	4	4	4
$P_1\ (P_A = A)$	18	20	25	30	42	63
$P_2\ (P_A = A)$	16	19	22	25	34	47
$P_3\ (P_A = A)$	5	6	8	8	8	8

Numerical Comparison (Flop Count)

Computational effort of GMRES and MINRES in megaflops.

n_z	4	8	16	32	64	128
N	94	314	1138	4322	16...	66434
K	2.73	93.44	$4K$	$191K$	$\star$	$\star$
$\tilde{K}$ ($P_A = A, P_H = H$)	0.03	0.25	2.90	38.52	548	$8,091$
$\tilde{K}$ ($P_A = A, P_H = I$)	0.04	0.27	2.61	30.77	418	$5,976$
P_1 ($P_A = A$)	0.12	0.73	6.41	64.48	779	$10,226$
P_2 ($P_A = A$)	0.11	0.71	5.94	58.66	708	$9,118$
P_3 ($P_A = A$)	0.05	0.42	4.33	44.86	528	$6,850$

Battermann/Heinkenschloss

- Mesh independence observed for all three preconditioners
- Slack variables from control constraints no problem
- Slack variable from state constraints introduce mesh dependence

In the latter case, the matrix

$$G = M_y^{1/2} \tilde{A}^{-1} B M_u^{-1/2}$$

has a large norm due to M_y.
Similar effect for small penalty parameters associated with M_u.

Outlook

- Exploit special structure of preconditioner 2
- Extend to cases with $L_{yu} \neq 0$
- Consider preconditioners $\tilde{M}_y, \tilde{M}_u$
- Scaling of slack variables
- Influence of penalty parameters
- Partially observed state variables, M_y not invertible