Preconditioning of Iterative Solvers for Partial Differential Equations

Andy Wathen
Computing Laboratory
Oxford University

wathen@comlab.ox.ac.uk
http://web.comlab.ox.ac.uk/~wathen/

Brief review of convergence results for Krylov subspace methods for $Ax = b$:

A symmetric and positive definite(SPD): Conjugate Gradients (CG) — iterates,

$$x_k \in x_0 + \underbrace{\mathrm{span}(r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0)}_{\mathcal{K}_k(A, r_0)}$$

which minimize $\|x - x_k\|_A$, hence convergence

$$\begin{aligned} \frac{\|x - x_k\|_A}{\|x - x_0\|_A} &\leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(A)} |p(\lambda)| \\ &\leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}. \end{aligned}$$

$$
\begin{aligned}
\frac{\|x - x_k\|_A}{\|x - x_0\|_A} &\leq \min_{p \in \Pi_k, p(0)=1} \; \max_{\lambda \in \sigma(A)} |p(\lambda)| \\
&\leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.
\end{aligned}
$$

Equivalently: takes

$$
k \approx \frac{1}{2} \log\left(\frac{\epsilon}{2}\right) \sqrt{\kappa}
$$

iterations to achieve

$$
\frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \epsilon
$$

$$CG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(A)} |p(\lambda)|$$

$$\leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

For fast convergence:

$$
\begin{aligned}
CG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} &\le \min_{p \in \Pi_k, p(0)=1} \; \max_{\lambda \in \sigma(A)} |p(\lambda)| \\
&\le 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.
\end{aligned}
$$

For fast convergence:

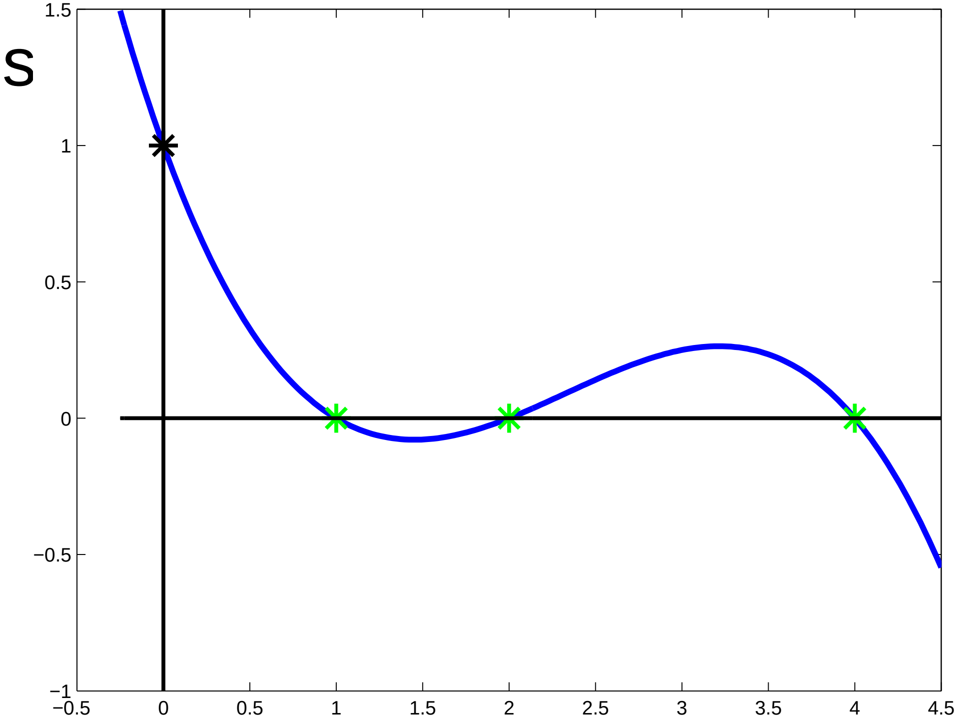
- few distinct eigenvalues
- cluster eigenvalues
- reduce κ

$$CG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \min_{p\in\Pi_k, p(0)=1} \max_{\lambda\in\sigma(A)} |p(\lambda)|$$

$$\leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

For fast convergence:

- few distinct eigenvalues
- cluster eigenvalues
- reduce κ

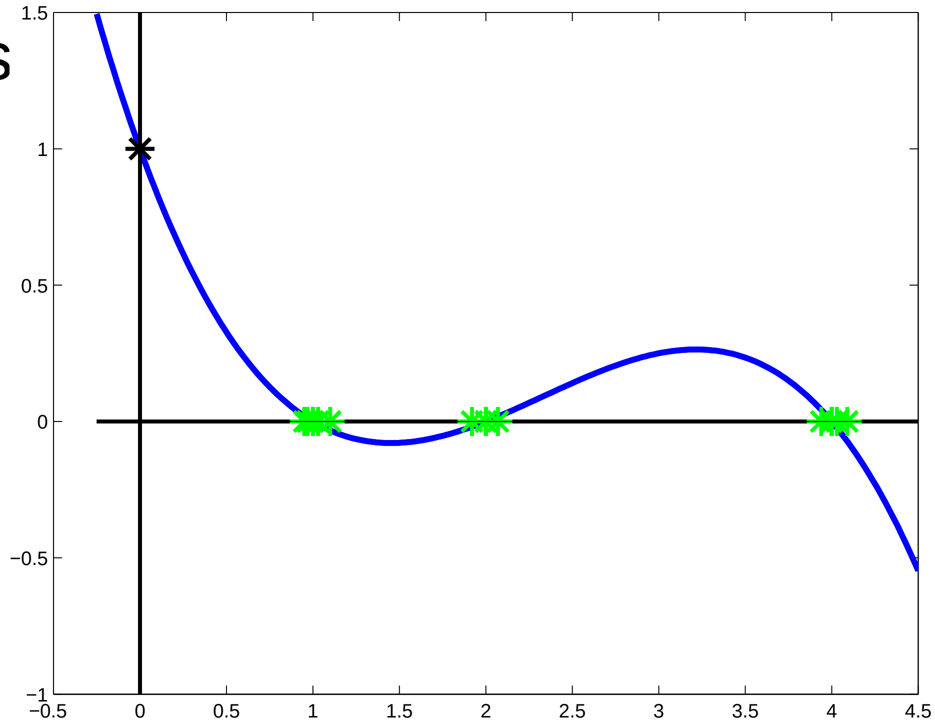


$$CG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(A)} |p(\lambda)|$$

$$\leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

For fast convergence:

- few distinct eigenvalues
- cluster eigenvalues
- reduce κ

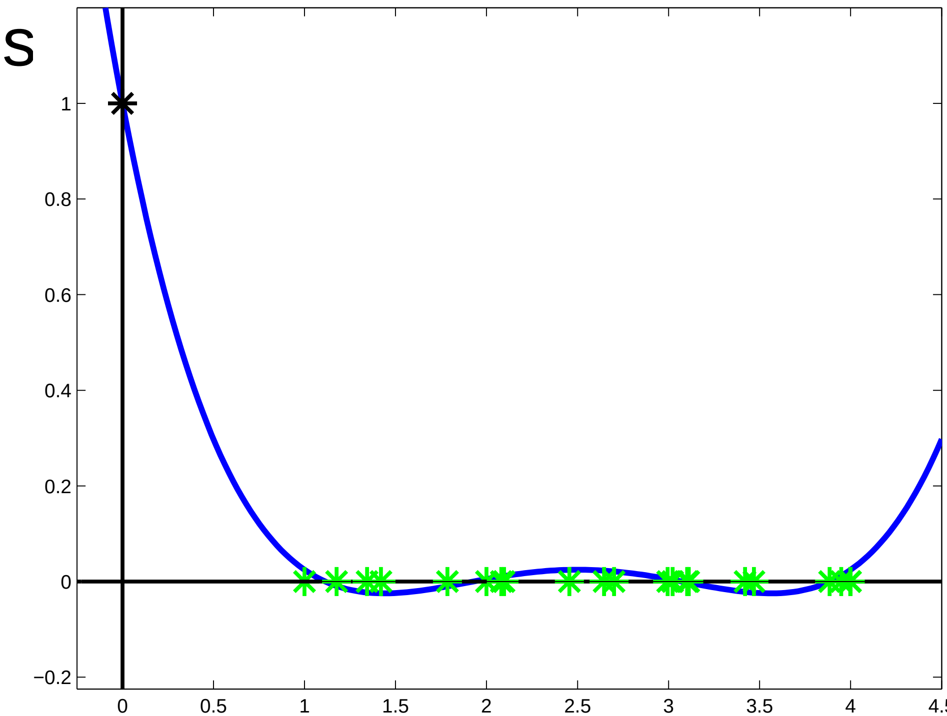


$$CG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(A)} |p(\lambda)|$$

$$\leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

For fast convergence:

- few distinct eigenvalues
- cluster eigenvalues
- reduce κ



Preconditioning:

matrix or linear operator $\boldsymbol{P}$ so that fast convergence (few eigenvalues/clustered eigenvalues/small $\boldsymbol{\kappa}$) for

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{P}^{-1}\boldsymbol{b}$$

in fact Preconditioned CG solves an equivalent symmetric matrix system

$$\boldsymbol{H}^{-T}\boldsymbol{A}\boldsymbol{H}^{-1}(\boldsymbol{H}\boldsymbol{x}) = \boldsymbol{H}^{-T}\boldsymbol{b}$$

where $\boldsymbol{P} = \boldsymbol{H}^T\boldsymbol{H}$ eg. $\boldsymbol{H}$ is Cholesky factor of $\boldsymbol{P}$ or $\boldsymbol{P}^{\frac{1}{2}}$.

$\Rightarrow$ preconditioner must be symmetric and is positive definite in all practical cases I know of!

Only $\boldsymbol{P}$ is needed, not $\boldsymbol{H}$: need to solve $\boldsymbol{P}\boldsymbol{z} = \boldsymbol{r}$ for $\boldsymbol{z}$ given $\boldsymbol{r}$ at each CG iteration

Note: $\kappa(A) = \lambda_{\max}(A)/\lambda_{\min}(A)$ happens to be 2-norm condition numbers associated with direct solution of linear system since A is symmetric.

$\kappa(H^{-T}AH^{-1}) = \frac{\lambda_{\max}(H^{-T}AH^{-1})}{\lambda_{\min}(H^{-T}AH^{-1})}$: because of similarity transform $H^{-1}(H^{-T}AH^{-1})H = P^{-1}A$ leads us to use

$$\kappa(P^{-1}A) = \frac{\lambda_{\max}(P^{-1}A)}{\lambda_{\min}(P^{-1}A)}$$

which is not the usual 2-norm condition number of the (nonsymmetric) matrix $P^{-1}A$

Note:

$$\frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} = \frac{\mathbf{v}^T H^{-T} A H^{-1} \mathbf{v}}{\mathbf{v}^T \mathbf{v}}, \quad \mathbf{u} = H^{-1}\mathbf{v}$$

Principle tool for analysis of preconditioning:

Generalised Rayleigh Quotient: because if

$$c \le \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} \le C$$

$$c \le \min_{\mathbf{u}} \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} = \lambda_{\min}(P^{-1}A) \quad ,$$

$$\lambda_{\max}(P^{-1}A) = \max_{\mathbf{u}} \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} \le C$$

$$\Rightarrow \quad \kappa(P^{-1}A) = \lambda_{\max}(P^{-1}A)/\lambda_{\min}(P^{-1}A) \le C/c$$

Hence preconditioned CG convergence

$$\frac{\|x - x_k\|_A}{\|x - x_0\|_A} \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \le 2\left(\frac{\sqrt{C}-\sqrt{c}}{\sqrt{C}+\sqrt{c}}\right)^k$$

Principle tool for analysis of preconditioning:

Generalised Rayleigh Quotient: because if

$$c \leq \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} \leq C$$

$$c \leq \min_{\mathbf{u}} \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} = \lambda_{\min}(P^{-1}A) \quad ,$$

$$\lambda_{\max}(P^{-1}A) = \max_{\mathbf{u}} \frac{\mathbf{u}^T A \mathbf{u}}{\mathbf{u}^T P \mathbf{u}} \leq C$$

$$\Rightarrow \quad \kappa(P^{-1}A) = \lambda_{\max}(P^{-1}A)/\lambda_{\min}(P^{-1}A) \leq C/c$$

Hence preconditioned CG convergence

$$\frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \leq 2\left(\frac{\sqrt{C}-\sqrt{c}}{\sqrt{C}+\sqrt{c}}\right)^k$$

Convergence norm unchanged for CG because:
convergence bound:

$$\|x - x_k\|_A \leq 2 \left(\frac{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}} \right)^k \|x - x_0\|_A$$

in $\|\cdot\|_A$ regardless of preconditioner, P, employed

Preconditioned CG $\leftrightarrow P^{-\frac{1}{2}} A P^{-\frac{1}{2}} y = P^{-\frac{1}{2}} b, \quad y = P^{\frac{1}{2}} x$
minimises

$$\|y - y_k\|_{P^{-\frac{1}{2}} A P^{-\frac{1}{2}}} = \|P^{\frac{1}{2}}(x - x_k)\|_{P^{-\frac{1}{2}} A P^{-\frac{1}{2}}} = \|x - x_k\|_A$$

over $x_k \in x_0 + \mathcal{K}_k(P^{-1}A, P^{-1}r_0), \quad r_0 = b - Ax_0$

In the context of PDEs':

usually not just 1 matrix but a family $A(h)$ of increasing dimension as h (the mesh size) gets smaller

(or $A(p)$ as the local polynomial degree, p, gets larger)

$\Rightarrow$ the dependence of κ on h is important:

If $\kappa = \mathcal{O}(h^{-t})$ takes $\sim h^{-\frac{t}{2}}$ iterations for convergence

eg. discrete Laplacian — (5-point/7-point) finite differences or finite elements: $\kappa = \mathcal{O}(h^{-2})$

$\Rightarrow$ on an $n \times n$ grid, $nh = \mathcal{O}(1)$, $A \in \mathrm{R}^{N \times N}, N = n^2$
CG takes $\sim n$ iterations for convergence

In the context of PDEs':

usually not just 1 matrix but a family $A(h)$ of increasing dimension as h (the mesh size) gets smaller

(or $A(p)$ as the local polynomial degree, p, gets larger)

$\Rightarrow$ the dependence of κ on h is important:

If $\kappa = \mathcal{O}(h^{-t})$ takes $\sim h^{-\frac{t}{2}}$ iterations for convergence

eg. discrete Laplacian — (5-point/7-point) finite differences or finite elements: $\kappa = \mathcal{O}(h^{-2})$

$\Rightarrow$ on an $n \times n$ grid, $nh = \mathcal{O}(1)$, $A \in \mathrm{R}^{N\times N}, N = n^2$
CG takes $\sim n$ iterations for convergence

Role of preconditioning: reduce h-dependence of κ

Ideal: make $\kappa(P^{-1}A)$ independent of h

In the context of PDEs':

usually not just 1 matrix but a family $A(h)$ of increasing dimension as h (the mesh size) gets smaller

(or $A(p)$ as the local polynomial degree, p, gets larger)

$\Rightarrow$ the dependence of κ on h is important:

If $\kappa = \mathcal{O}(h^{-t})$ takes $\sim h^{-\frac{t}{2}}$ iterations for convergence

eg. discrete Laplacian — (5-point/7-point) finite differences or finite elements: $\kappa = \mathcal{O}(h^{-2})$

$\Rightarrow$ on an $n \times n$ grid, $nh = \mathcal{O}(1)$, $A \in \mathrm{R}^{N \times N}, N = n^2$
CG takes $\sim n$ iterations for convergence

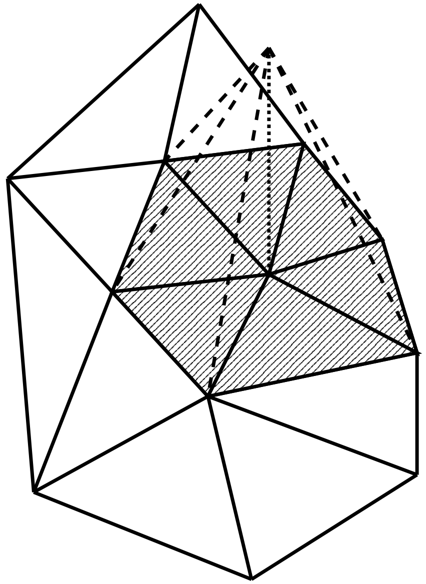
Role of preconditioning: reduce h-dependence of κ

Ideal: make $\kappa(P^{-1}A)$ independent of h

(then P and A are Spectrally Equivalent)

SPD Example 1: finite element mass matrix (conforming approximation)

$\boldsymbol{u_h(\mathrm{x}) = \sum_{j=1}^{N} \mathrm{u}_j \phi_j(\mathrm{x}) \in X_h \subset \mathcal{H}^1(\Omega)}$



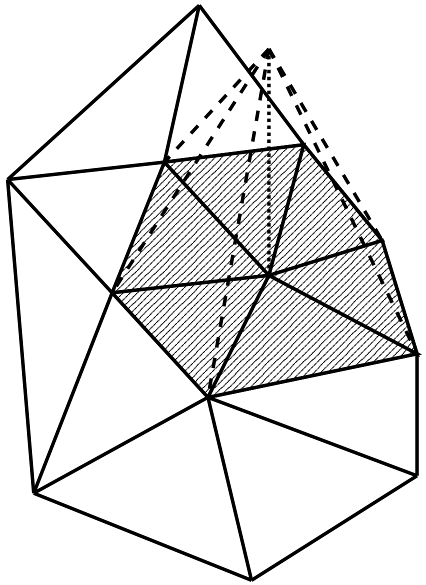
Mass matrix is 'identity operator':
(or $\boldsymbol{L_2(\Omega)}$ projection)

$$\boldsymbol{\|u_h\|^2 = \sum_{j=1}^{N}\sum_{i=1}^{N} \mathrm{u}_i \mathrm{u}_j \int_\Omega \phi_i \phi_j \mathrm{d}\Omega = \mathrm{u}^T Q \mathrm{u}}$$

$\boldsymbol{Q = \{q_{i,j}, i, j = 1, \ldots, N\}, \;\; q_{i,j} = \int_\Omega \phi_i \phi_j}$

SPD Example 1: finite element mass matrix (conforming approximation)

$u_h(\mathrm{x}) = \sum_{j=1}^{N} \mathbf{u}_j \phi_j(\mathrm{x}) \in X_h \subset \mathcal{H}^1(\Omega)$



Mass matrix is 'identity operator':
(or $L_2(\Omega)$ projection)

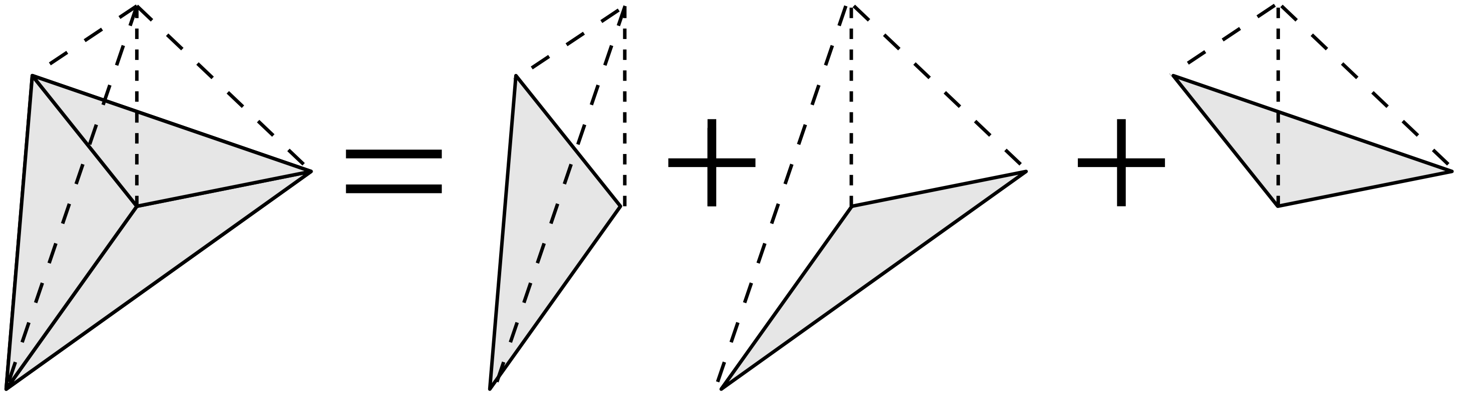
$$\|u_h\|^2 = \sum_{j=1}^{N} \sum_{i=1}^{N} \mathbf{u}_i \mathbf{u}_j \int_\Omega \phi_i \phi_j \mathrm{d}\Omega = \mathbf{u}^T Q \mathbf{u}$$

$Q = \{q_{i,j}, i, j = 1, \ldots, N\}, \; q_{i,j} = \int_\Omega \phi_i \phi_j$

sparse, but not the identity matrix!

SPD Example 1: finite element mass matrix

Assembled from element contributions



Mathematically: if $Q_e \in \mathrm{R}^{\nu \times \nu}$ is the element mass matrix for element $e = 1, 2, \ldots, E$ then

$$Q = L^T [Q_e] L$$

where $[Q_e] \equiv \mathbf{diag}(Q_1, Q_2, \ldots, Q_E)$ and $L \in \mathrm{R}^{E\nu \times N}$ is the connectivity matrix with 1's in column j for each row where a local basis function on an element is part of the j^{th} global basis function ϕ_j

SPD Example 1: finite element mass matrix

Observation: as well as

$$\boldsymbol{Q} = \boldsymbol{L}^{\boldsymbol{T}}[\boldsymbol{Q}_{\boldsymbol{e}}]\boldsymbol{L}$$

it is also always true that

$$\mathbf{diag}(\boldsymbol{Q}) = \boldsymbol{L}^{\boldsymbol{T}}[\mathbf{diag}(\boldsymbol{Q}_{\boldsymbol{e}})]\boldsymbol{L}$$

ie. assembly of the diagonal elements of the element matrices is exactly the diagonal of the global matrix (*Freid, W*)

Hence eigenvalues $\boldsymbol{\lambda}$ of $\mathbf{diag}(Q)^{-1}Q$ all satisfy

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}}$$

Hence eigenvalues $\boldsymbol{\lambda}$ of $\mathbf{diag}(Q)^{-1}Q$ all satisfy

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}}$$

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}}$$

Hence eigenvalues $\boldsymbol{\lambda}$ of $\mathbf{diag}(Q)^{-1}Q$ all satisfy

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}}$$

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}}$$

$$\min_{\mathbf{y}=L\mathbf{x}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{y}=L\mathbf{x}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}}$$

Hence eigenvalues $\boldsymbol{\lambda}$ of $\mathbf{diag}(Q)^{-1}Q$ all satisfy

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{diag}(Q) \mathbf{x}}$$

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{x}} \frac{\mathbf{x}^T L^T [Q_e] L \mathbf{x}}{\mathbf{x}^T L^T [\mathbf{diag}(Q_e)] L \mathbf{x}}$$

$$\min_{\mathbf{y}=L\mathbf{x}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{y}=L\mathbf{x}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}}$$

$$\Rightarrow \min_{\mathbf{y}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}} \leq \boldsymbol{\lambda} \leq \max_{\mathbf{y}} \frac{\mathbf{y}^T [Q_e] \mathbf{y}}{\mathbf{y}^T [\mathbf{diag}(Q_e)] \mathbf{y}}$$

ie. the extreme eigenvalues of the diagonally preconditioned (scaled) element matrices bound the extreme eigenvalues of the diagonally preconditioned (scaled) global matrix.

Example: P1 (piecewise linear) trianglular elements

$$Q_e = \frac{1}{6}\ulcorner_e \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \ulcorner_e = \text{area of element } e.$$

Thus for every element e

$$\mathrm{diag}(Q_e)^{-1} Q_e = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

has eigenvalues $2, \frac{1}{2}, \frac{1}{2} \Rightarrow$ for every eigenvalue of the preconditioned global mass matrix satisfies

$$\frac{1}{2} \le \lambda(\mathrm{diag}(Q)^{-1} Q) \le 2 \quad \Rightarrow \kappa = \lambda_{\max}/\lambda_{\min} \le 4 \text{ and}$$

$$PCG: \quad \frac{\|x - x_k\|_A}{\|x - x_0\|_A} \le 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \le 2(\frac{1}{3})^k$$

SPD Example 1: finite element mass matrix

in fact without preconditioning eigenvalues of Q are independent of h but dependent on element sizes and any variable coefficients (density) which simple diagonal preconditioning (also called Jacobi preconditioning) removes.

SPD Example 1: finite element mass matrix

in fact without preconditioning eigenvalues of Q are independent of h but dependent on element sizes and any variable coefficients (density) which simple diagonal preconditioning (also called Jacobi preconditioning) removes.

demo: 'serendipity' (8-node) finite element for which

$$\frac{1}{4} \le \lambda(\text{diag}(Q)^{-1}Q) \le \frac{9}{2}$$

$\Rightarrow \kappa \le 18$

$$\Rightarrow PCG: \frac{\|x - x_k\|_Q}{\|x - x_0\|_Q} \le 2(0.6185)^k$$

SPD Example 2: Discrete Laplacian (Dirichlet problem)

$$-\nabla^2 u = f \quad \text{in } \Omega, \quad u = g \text{ on } \partial\Omega$$

Finite Differences:

$$A \sim h^{-2} \quad \begin{array}{ccccc} & & -1 & & \\ & & | & & \\ -1 & — & 4 & — & -1 \\ & & | & & \\ & & -1 & & \end{array}$$

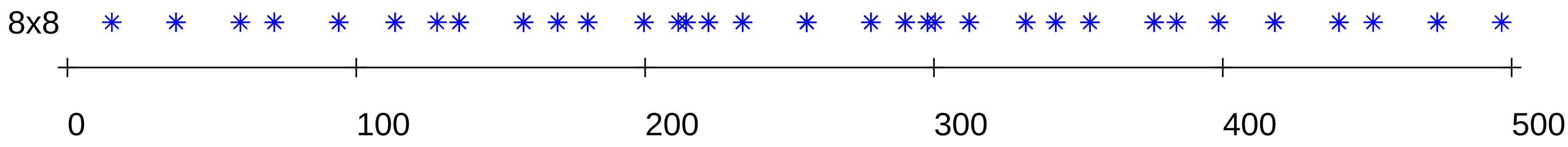
eg. on unit square A is block tridiagonal:

$$\underbrace{h^{-2}\begin{bmatrix} B & -I & & & \\ -I & B & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & B & -I \\ & & 0 & -I & B \end{bmatrix}}_{A \in \mathbb{R}^{n^2 \times n^2}}, \quad \underbrace{\begin{bmatrix} 4 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 4 \end{bmatrix}}_{B \in \mathbb{R}^{n \times n}}$$

5-point finite difference discrete Laplacian
Using discrete Fourier analysis eigenvalues known:

$$\lambda = h^{-2}\left[4 - 2\cos(r\pi h) - 2\cos(s\pi h)\right], \quad r, s = 1, \ldots, n$$

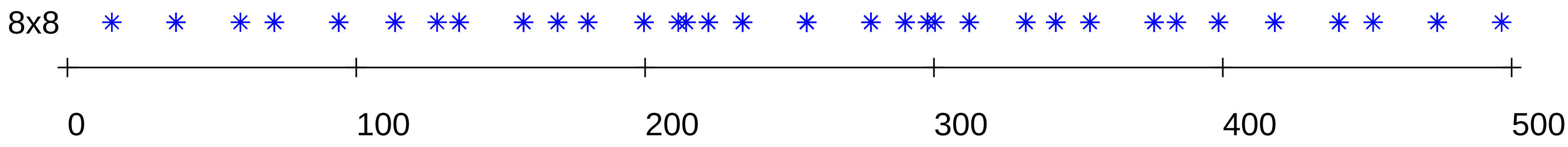
$$\Rightarrow \qquad \kappa = \frac{4}{\pi^2}\, h^{-2}, \qquad \lambda_{\min} \approx 2\pi^2,\ \lambda_{\max} \approx 8h^{-2}$$

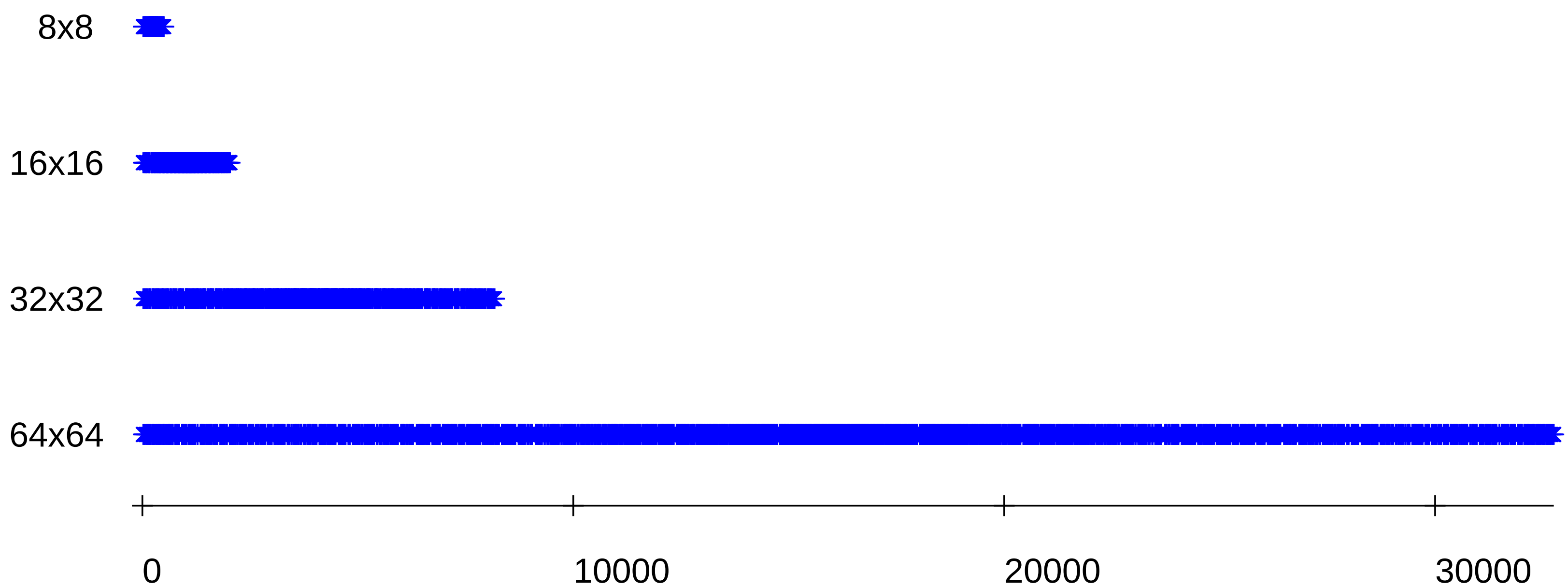


5-point finite difference discrete Laplacian
Using discrete Fourier analysis eigenvalues known:

$$\lambda = h^{-2}\,[4 - 2\cos(r\pi h) - 2\cos(s\pi h)]\,, \quad r, s = 1, \ldots, n$$

$$\Rightarrow \qquad \kappa = \frac{4}{\pi^2}\,h^{-2}, \qquad \lambda_{\min} \approx 2\pi^2,\ \lambda_{\max} \approx 8h^{-2}$$

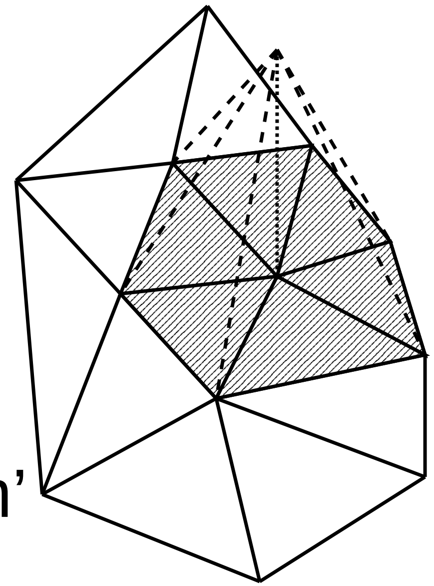




SPD Example 2: Discrete Laplacian (Dirichlet problem)

Finite Elements - Stiffness matrix:

$\boldsymbol{u_h(\mathbf{x}) = \sum_{j=1}^{N} \mathbf{u}_j \phi_j(\mathbf{x}) \in X_h \subset \mathcal{H}^1(\Omega)}$



Stiffness matrix: Grammian in ‘energy norm’

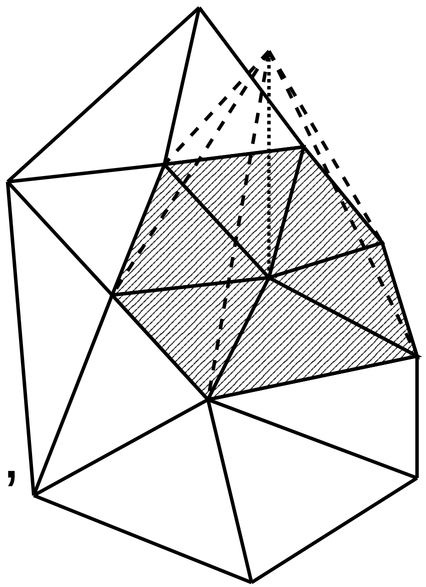
$$\|\nabla u_h\|^2 = \sum_{j=1}^{N} \sum_{i=1}^{N} \mathbf{u}_i \mathbf{u}_j \int_\Omega \nabla \phi_i \cdot \nabla \phi_j \mathbf{d}\Omega = \mathbf{u}^T A \mathbf{u}$$

$\boldsymbol{A = \{a_{i,j}, i, j = 1, \ldots, N\}, \;\; a_{i,j} = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j}$

SPD Example 2: Discrete Laplacian (Dirichlet problem)

Finite Elements - Stiffness matrix:

$u_h(\mathbf{x}) = \sum_{j=1}^{N} \mathbf{u}_j \phi_j(\mathbf{x}) \in X_h \subset \mathcal{H}^1(\Omega)$



Stiffness matrix: Grammian in ‘energy norm’

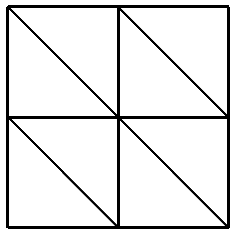
$$\|\nabla u_h\|^2 = \sum_{j=1}^{N} \sum_{i=1}^{N} \mathbf{u}_i \mathbf{u}_j \int_\Omega \nabla \phi_i \cdot \nabla \phi_j \mathrm{d}\Omega = \mathbf{u}^T A \mathbf{u}$$

$A = \{a_{i,j}, i, j = 1, \ldots, N\}, \;\; a_{i,j} = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j$

sparse: the target of much research!

SPD Example 2: Discrete Laplacian (Dirichlet problem) - Stiffness matrix

Finite Elements: in fact for P1 (piecewise linear on triangles) arranged as



exactly same matrix A as 5-point finite differences

for Q1 (bilinear on square elements)

$$A \sim \begin{array}{ccccc} -1 & & -1 & & -1 \\ & \diagdown & | & \diagup & \\ -1 & - & 8 & - & -1 \\ & \diagup & | & \diagdown & \\ -1 & & -1 & & -1 \end{array}$$

Preconditioners:

- Incomplete Factorization
- Direct Sparse methods/Support Graph preconditioning
- Sparse Approximate Inverses
- Domain Decomposition
- Multigrid/Multilevel preconditioning

Incomplete Factorization: factorization maintaining sparsity

- Incomplete orthogonal (QR) factorization (some papers: limited usefulness for PDEs)
- Incomplete triangular (LU or LL^T) factorization: (*Meijerink & van der Vorst*)

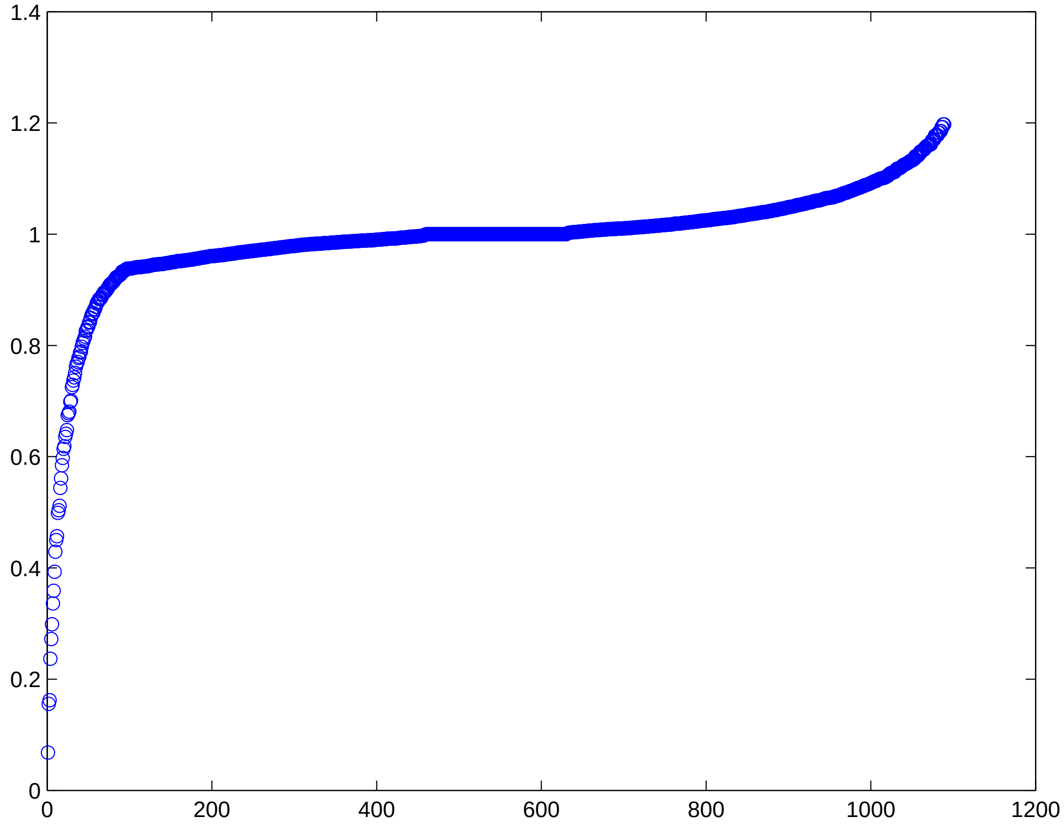
$$
\begin{array}{l}
\texttt{for } i = 1 \texttt{ until } n \texttt{ do} \\
\quad m = \min\{k : a_{ik} \neq 0\} \\
\quad \texttt{for } j = m \texttt{ until } i-1 \texttt{ do} \\
\quad\quad \texttt{if } a_{ij} \neq 0 \texttt{ then} \\
\quad\quad\quad l_{ij} \leftarrow a_{ij} - \sum_{k=m}^{j-1} l_{ik} l_{jk} / l_{jj} \\
\quad\quad \texttt{endif} \\
\quad \texttt{enddo} \\
\quad l_{ii} = \left(a_{ii} - \sum_{k=m}^{i-1} l_{ik} l_{ik} \right)^{1/2} \\
\texttt{enddo}
\end{array}
$$

Many variants: IC (cholinc), ILU (luinc), MIC, MILU, ILUT, ILUTP,...

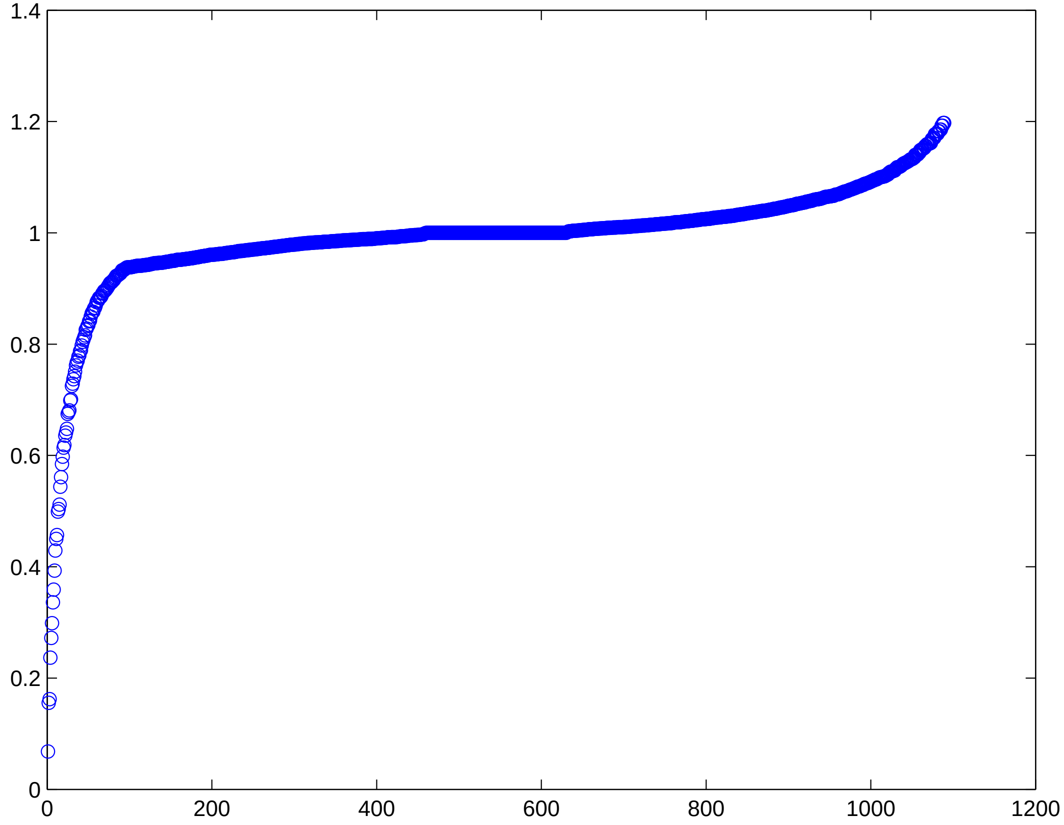
Theory: mostly for M-matrices and generalisations, in particular model Poisson problem:

- IC: $\boldsymbol{\lambda}_{\min} = \mathcal{O}(h), \;\; \boldsymbol{\lambda}_{\max} = \mathcal{O}(h^{-1})$ for 2D Poisson $\Rightarrow \boldsymbol{\kappa} = \mathcal{O}(h^{-2})$
- MIC: diagonal modified to preserve row sums (*Gustafsson*): $\boldsymbol{\lambda}_{\min} = \mathcal{O}(1), \;\; \boldsymbol{\lambda}_{\max} = \mathcal{O}(h^{-1})$ for 2D Poisson $\Rightarrow \boldsymbol{\kappa} = \mathcal{O}(h^{-1})$

Incomplete Cholesky - Practice IC(0) for Q1 finite elements:



Incomplete Cholesky - Practice IC(0) for Q1 finite elements:



- Advantage: algebraic
- Disadvantage: does not scale for PDEs

Direct Sparse elimination:

after ‘sparsification’ or for related problem can be used as a preconditioner

Support Graph preconditioning:

based on sparse elimination data structures (modified spanning tree): related to ILU

(*Vaidya, Boman & Hendrickson*)

Sparse Approximate Inverses: select desired sparsity pattern $\mathcal{S}$ and

$$\min_{P \in \mathcal{S}} \|I - AP\|_F$$

separates into n independent linear least squares problems for columns of P (in parallel). Variants:

- in factored form (symmetry) (*Kolotilina & Yeremin, Axelsson*)
- dynamic sparsity selection: SPAI (*Huckle & Grote*)
- different norms (*Axelsson, Holland, Shaw & W*)
- ‘target’ matrices: $\min_{P \in \mathcal{S}} \|T - AP\|_F$ (*Holland*)
- AINV (*Benzi & Tuma*)

Sparse Approximate Inverses:

- Advantage: algebraic
- Disadvantage: does not apparently scale for PDEs

also provide an interesting class of ‘smoothers’ for multigrid (*Tang & Wan*)

Domain Decomposition:

split Ω into domains Ω_i such that $\cup_i \Omega_i = \Omega$.
Use solvers on each subdomain as preconditioner for whole problem $\leftrightarrow$ Block Diagonal Preconditioning.

Issues:

- overlaping or non-overlaping subdomains
- boundary conditions on subdomains
- coarse grid problem
- PARALLEL COMPUTING

Very widely researched: books of Smith, Bjørstad & Gropp, Quarteroni & Valli. also software: PETSc,...

Domain Decomposition:

Typically if $|\Omega_i| = H$ then

$$\kappa \sim H^2 \left(1 + \log \frac{H}{h}\right)^2$$

without coarse grid solve, or

$$\kappa \sim \left(1 + \log \frac{H}{h}\right)^2$$

with a coarse grid solve

Multigrid/Multilevel preconditioning $\longrightarrow$ Yvan Notay

appropriate methods (smoothing and grid transfers) converge in a number of iterations independent of h
$\Rightarrow$ optimal solvers for Poisson problems

Number of V-cycles (contraction factor: η)

$\|\mathbf{r}^{(k)}\|/\|\mathbf{r}^{(0)}\| \leq 10^{-4}$ for Q_1 (bilinear) elements

number damped Jacobi pre- and post-smoothing is ℓ–m.

grid	1–0	1–1	2–0	n
8×8	7 (0.25)	4 (0.10)	4 (0.09)	49
16×16	13 (0.43)	4 (0.11)	5 (0.11)	225
32×32	20 (0.57)	4 (0.12)	6 (0.14)	961
64×64	31 (0.66)	4 (0.14)	7 (0.15)	3969
128×128	45 (0.73)	5 (0.16)	8 (0.15)	16129

Multigrid:

Convergence bound: $\|\mathbf{u}-\mathbf{u}^{(k)}\|_A \leq \eta\|\mathbf{u}-\mathbf{u}^{(k-1)}\|_A$
η typically 0.1

$\Rightarrow$ multigrid is a great preconditioner for Laplacian because

$$\|\mathbf{u}-\mathbf{u}^{(k)}\|_A \leq \eta\|\mathbf{u}-\mathbf{u}^{(k-1)}\|_A$$

$$\Rightarrow 1-\eta \leq \lambda_{\min}(P^{-1}A), \lambda_{\max}(P^{-1}A) \leq 1+\eta$$

when P^{-1} is the action of a single multigrid cycle.
Hence $\kappa \leq (1+\eta)/(1-\eta) = \mathcal{O}(1)$

Multigrid:

Convergence bound: $\|\mathbf{u} - \mathbf{u}^{(k)}\|_A \leq \eta \|\mathbf{u} - \mathbf{u}^{(k-1)}\|_A$
η typically 0.1

$\Rightarrow$ multigrid is a great preconditioner for Laplacian because

$$\|\mathbf{u} - \mathbf{u}^{(k)}\|_A \leq \eta \|\mathbf{u} - \mathbf{u}^{(k-1)}\|_A$$

$$\Rightarrow 1 - \eta \leq \lambda_{\min}(P^{-1}A), \lambda_{\max}(P^{-1}A) \leq 1 + \eta$$

when P^{-1} is the action of a single multigrid cycle.
Hence $\kappa \leq (1+\eta)/(1-\eta) = \mathcal{O}(1)$

Algebraic versions: AMG (*Ruge & Stüben, Henson*)

Comment: ‘right’ norm for convergence with finite elements

$u \in X$: a function we want to find

Approximation:

$u_h = \sum_j \mathbf{u}_j \phi_j \in X_h \subset X$: a function we will compute

Vector of coefficients $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)^T$ will be discrete (nodal) values of the approximate function (solution) when $\phi_j(\mathbf{x}_i) = \delta_{i,j}$

Let $\mathbf{u}^{(k)} = (\mathbf{u}_1^{(k)}, \mathbf{u}_2^{(k)}, \ldots, \mathbf{u}_n^{(k)})^T$ be k^{th} iterate for solution of $A\mathbf{u} = \mathbf{f}$

Useful point: $\mathbf{u}^{(k)} \leftrightarrow u_h^{(k)}$ via $u_h^{(k)} = \sum_j \mathbf{u}_j^{(k)} \phi_j$

and in any norm

$$\|u - u_h^{(k)}\| \leq \underbrace{\|u - u_h\|}_{\text{FE error}} + \underbrace{\|u_h - u_h^{(k)}\|}_{\text{iteration error}}$$

$$\|u - u_h^{(k)}\| \leq \underbrace{\|u - u_h\|}_{\text{FE error}} + \underbrace{\|u_h - u_h^{(k)}\|}_{\text{iteration error}}$$

FE error: (a priori) error estimate,
iteration error: iterative convergence bound

In particular for Galerkin finite elements for the Laplacian we have seen

$$\|\nabla v_h\|^2 = \|\mathbf{v}\|_A^2$$

so that

$$\|\nabla(u - u_h^{(k)})\| = \underbrace{\|\nabla(u - u_h)\|}_{\text{FE error}} + \underbrace{\|\mathbf{u} - \mathbf{u}^{(k)}\|_A}_{\text{iteration error}}$$

$\Rightarrow$ regardless of preconditioner used we get CG (and MG) convergence in exactly the correct norm: can ensure that $\|\mathbf{u} - \mathbf{u}^{(k)}\|_A$ is comparable to $\|\nabla(u - u_h)\|$

Multigrid:

$\|\mathbf{r}^{(k)}\|/\|\mathbf{r}^{(0)}\| \leq 10^{-4}h^2$ for Q_1 (bilinear) elements

number damped Jacobi pre- and post-smoothing is ℓ–m.

grid	1–0	1–1	2–0	$10^{-4}\,h^2$
8×8	9	5	5	1.6×10^{-6}
16×16	17	6	7	3.9×10^{-7}
32×32	30	7	9	9.8×10^{-8}
64×64	48	8	11	2.4×10^{-8}
128×128	71	9	13	6.1×10^{-9}

Multigrid:

$\|\mathbf{r}^{(k)}\|/\|\mathbf{r}^{(0)}\| \leq \mathbf{10}^{-4}\boldsymbol{h^2}$ for $\boldsymbol{Q_1}$ (bilinear) elements

number damped Jacobi pre- and post-smoothing is $\boldsymbol{\ell}$–$\boldsymbol{m}$.

grid	1–0	1–1	2–0	$\mathbf{10}^{-4}\,\boldsymbol{h^2}$
$\mathbf{8 \times 8}$	9	5	5	$\mathbf{1.6 \times 10^{-6}}$
$\mathbf{16 \times 16}$	17	6	7	$\mathbf{3.9 \times 10^{-7}}$
$\mathbf{32 \times 32}$	30	7	9	$\mathbf{9.8 \times 10^{-8}}$
$\mathbf{64 \times 64}$	48	8	11	$\mathbf{2.4 \times 10^{-8}}$
$\mathbf{128 \times 128}$	71	9	13	$\mathbf{6.1 \times 10^{-9}}$

(so called Full Multigrid cycle can remove the slight iteration growth)

Other SPD problems:

- variable coefficients: $-\boldsymbol{\nabla} \cdot (\boldsymbol{K\nabla u}) = \boldsymbol{f}$
- higher dimensions (finance)
- higher order (Biharmonic)

Other SPD problems:

- variable coefficients: $-\boldsymbol{\nabla} \cdot (\boldsymbol{K}\boldsymbol{\nabla}\boldsymbol{u}) = \boldsymbol{f}$
- higher dimensions (finance)
- higher order (Biharmonic)

demo: Poisson (diffusion) problem with IFISS

Symmetric indefinite problems (SID):

iterative methods: MINRES (also SYMMLQ) (*Paige & Saunders*)
(CG is not robust)

For $\mathcal{A}\mathbf{u} = \mathbf{f}$, MINRES minimises $\|\mathbf{r}^{(k)}\|$, $\mathbf{r}^{(k)} = \mathbf{f} - \mathcal{A}\mathbf{u}^{(k)}$

Convergence:

$$\frac{\|\mathbf{r}^{(k)}\|_{\mathcal{A}}}{\|\mathbf{r}^{(0)}\|_{\mathcal{A}}} \leq \min_{p\in\Pi_k, p(0)=1} \max_{\lambda\in\sigma(\mathcal{A})} |p(\lambda)|$$

$\Rightarrow$ if $\lambda \in [-a, -b] \cup [c, d]$, $0 < a, b, c, d$ then convergence independent of h if a, b, c, d are independent of h
Convergence bounds more complicated with h-dependence

For some problems $\kappa = bc/ad$ plays a similar role to κ for CG for SPD problems

SID Example 1: Helmholtz problem

$$-\nabla^2 u - ku = f$$

is a difficult problem for iteration since it has resonances (zero eigenvalues) for certain k, h

Multigrid possible: *Elman, Osterlee*

SID Example 1: Helmholtz problem

$$-\nabla^2 u - k u = f$$

is a difficult problem for iteration since it has resonances (zero eigenvalues) for certain k, h

Multigrid possible: *Elman, Osterlee*

SID Example 2: Stokes problem

$$\begin{aligned} -\nabla^2 u + \nabla p &= f \\ \nabla \cdot u &= 0 \end{aligned}$$

SID Example 2: Stokes problem

$$\begin{aligned} -\nabla^2 u + \nabla p &= f \\ \nabla \cdot u &= 0 \end{aligned}$$

Mixed finite element approximation:

$$u \approx u_h = \sum_{j=1}^{N} \mathrm{u}_j \phi_j(\mathrm{x}) \in X_h \subset \mathcal{H}^1$$
$$p \approx p_h = \sum_{\ell=1}^{M} \mathrm{p}_\ell \psi_\ell(\mathrm{x}) \in M_h \subset L_2$$

$$\Rightarrow \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}$$

A: discrete Laplacian for each component of the velocity

$B = \{B_{\ell,j}, b_{\ell,j} = \int_\Omega \psi_\ell \nabla \cdot \phi_j\}$: discrete divergence

$\leftrightarrow B^T$: discrete gradient

B generally full rank $\leftrightarrow$ inf-sup (LBB) stability

Stokes:

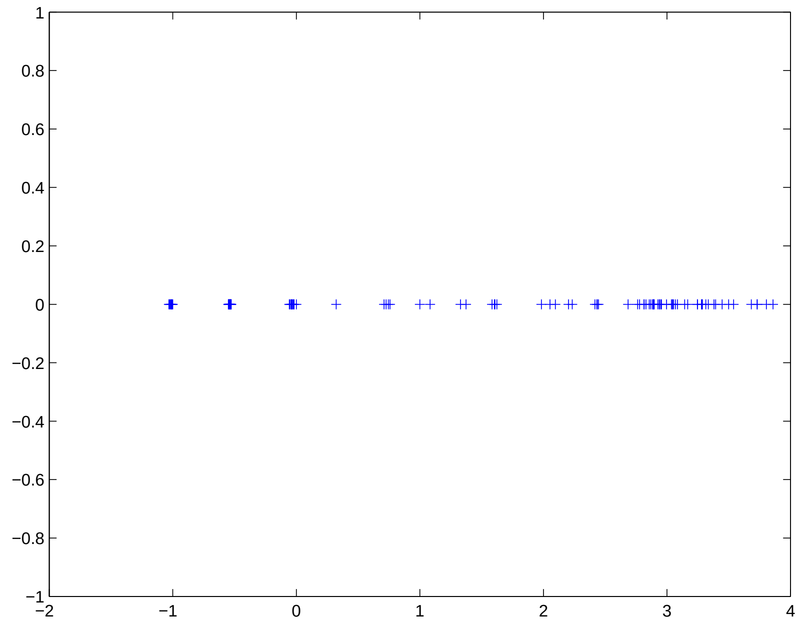
$$\begin{aligned} -\nabla^2 \boldsymbol{u} + \nabla \boldsymbol{p} &= \boldsymbol{f} \\ \nabla \cdot \boldsymbol{u} &= \mathbf{0} \end{aligned}$$

$\Rightarrow$

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B}^T \\ \boldsymbol{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

Symmetric Indefinite:

eigenvalues:



SID Example 2: Stokes problem

Indefinite: congruence

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & -BA^{-1}B^T \end{bmatrix} \begin{bmatrix} I & A^{-1}B^T \\ 0 & I \end{bmatrix}$$

(Typical) error estimate:

$$\begin{aligned} \|\nabla(u-u_h)\| &+ \|p-p_h\| \\ &\leq C(\inf_{v_h \in X_h} \|\nabla(u-v_h)\| + \inf_{q_h \in M_h} \|p-q_h\|) \\ &\leq C\,h^2 \left(\|D^3 u\| + \|D^2 p\|\right), \end{aligned}$$

$\Rightarrow$ prefered iterative convergence in

$$\|\mathrm{u}-\mathrm{u}^{(k)}\|_A + \|\mathrm{p}-\mathrm{p}^{(k)}\|_Q$$

Stokes

$$\mathcal{A}\mathbf{x} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix}$$ $\rightarrow$ symmetric indefnite $\Rightarrow$ MINRES

MINRES minimises $\|$residual$\|_I$
and preconditioned MINRES minimises $\|$residual$\|_{P^{-1}}$,
P: preconditioner

Preconditioned MINRES
$\leftrightarrow P^{-\frac{1}{2}}\mathcal{A}P^{-\frac{1}{2}}\mathbf{y} = P^{-\frac{1}{2}}\mathbf{f}, \quad \mathbf{y} = P^{\frac{1}{2}}\mathbf{x}$
minimises

$$\|P^{-\frac{1}{2}}\mathbf{f} - P^{-\frac{1}{2}}\mathcal{A}P^{-\frac{1}{2}}\mathbf{y})\|_I = \|P^{-\frac{1}{2}}(\mathbf{f} - \mathcal{A}\mathbf{x})\|_I = \|\mathbf{f} - \mathcal{A}\mathbf{x}\|_{P^{-1}}$$

over $\mathbf{x}^{(k)} \in \mathbf{x}^{(0)} + \mathcal{K}(P^{-1}\mathcal{A}, P^{-1}\mathbf{r}^{(0)}), \quad \mathbf{r}^{(0)} = \mathbf{f} - \mathcal{A}\mathbf{x}^{(0)}$

Note: SPD preconditioner needed for SID problem

Stokes

$\boldsymbol{\mathcal{A}}\mathbf{x} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B}^T \\ \boldsymbol{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix}$ $\longrightarrow$ symmetric indefnite $\Rightarrow$ MINRES

MINRES minimises $\|$residual$\|_{\boldsymbol{I}}$

and preconditioned MINRES minimises $\|$residual$\|_{\boldsymbol{P}^{-1}}$,

$\boldsymbol{P}$: preconditioner

Recall error for Stokes $\|\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)\| + \|\boldsymbol{p} - \boldsymbol{p}_h\| \Rightarrow$ prefered iterative convergence in

$\|\mathbf{u} - \mathbf{u}^{(k)}\|_{\boldsymbol{A}} + \|\mathbf{p} - \mathbf{p}^{(k)}\|_{\boldsymbol{Q}}$

$\Leftrightarrow$ $\|$error$\|_{\boldsymbol{E}}$ $=$ $\|$residual$\|_{\boldsymbol{\mathcal{A}}^{-1}\boldsymbol{E}\boldsymbol{\mathcal{A}}^{-1}}$ $=$ $\|$residual$\|_{\boldsymbol{H}^{-1}}$

where

$$\boldsymbol{E} = \begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q} \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} \boldsymbol{A} + \boldsymbol{B}^T\boldsymbol{Q}^{-1}\boldsymbol{B} & \boldsymbol{B}^T \\ \boldsymbol{B} & \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T \end{bmatrix}$$

Stokes

$$\mathcal{A}\mathbf{x} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix}$$ → symmetric indefnite ⇒ MINRES

MINRES minimises $\|$residual$\|_I$

and preconditioned MINRES minimises $\|$residual$\|_{P^{-1}}$,

P: preconditioner

Recall error for Stokes $\|\nabla(u - u_h)\| + \|p - p_h\| \Rightarrow$

prefered iterative convergence in

$\|\mathbf{u} - \mathbf{u}^{(k)}\|_A + \|\mathbf{p} - \mathbf{p}^{(k)}\|_Q$

$$\Leftrightarrow \|\text{error}\|_E = \|\text{residual}\|_{\mathcal{A}^{-1}E\mathcal{A}^{-1}} = \|\text{residual}\|_{H^{-1}}$$

where

$$E = \begin{bmatrix} A & 0 \\ 0 & Q \end{bmatrix}, \quad H = \begin{bmatrix} A + B^T Q^{-1} B & B^T \\ B & BA^{-1}B^T \end{bmatrix}$$

⇒ H as preconditioner!!

Stokes: Summarize: Taking preconditioner $\boldsymbol{H}$ with MINRES will give iterative convergence in ‘right’ norm

not practical(!)

Stokes: Summarize: Taking preconditioner $\boldsymbol{H}$ with MINRES will give iterative convergence in ‘right’ norm

not practical(!) but the block diagonal preconditioner

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{M} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T} \end{bmatrix}$$

is equivalent to $\boldsymbol{H}$ in the sence that

$$\boldsymbol{c}\|\text{error}\|_{\boldsymbol{E}} \leq \|\text{residual}\|_{\boldsymbol{P}^{-1}} \leq \boldsymbol{C}\|\text{error}\|_{\boldsymbol{E}}$$

whenever $\boldsymbol{M} \sim \boldsymbol{A}$ and $\boldsymbol{T} \sim \boldsymbol{Q}$ are spectral equivalences ie. whenever
$\boldsymbol{M}$ as preconditioner for $\boldsymbol{A}$ would give $\boldsymbol{h}$-independent CG convergence

and $\boldsymbol{T}$ as preconditioner for $\boldsymbol{Q}$ would give $\boldsymbol{h}$-independent CG

convergence

Stokes: Summarize: Taking preconditioner $\boldsymbol{H}$ with MINRES will give iterative convergence in ‘right’ norm

not practical(!) but the block diagonal preconditioner

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{M} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T} \end{bmatrix}$$

is equivalent to $\boldsymbol{H}$ in the sence that

$$\boldsymbol{c}\|\text{error}\|_{\boldsymbol{E}} \leq \|\text{residual}\|_{\boldsymbol{P}^{-1}} \leq \boldsymbol{C}\|\text{error}\|_{\boldsymbol{E}}$$

whenever $\boldsymbol{M} \sim \boldsymbol{A}$ and $\boldsymbol{T} \sim \boldsymbol{Q}$ are spectral equivalences ie. whenever
$\boldsymbol{M}$ as preconditioner for $\boldsymbol{A}$ would give $\boldsymbol{h}$-independent CG converg ence $\Rightarrow$ use $\boldsymbol{M}$ is a multigrid cycle for Laplacian
and $\boldsymbol{T}$ as preconditioner for $\boldsymbol{Q}$ would give $\boldsymbol{h}$-independent CG convergence $\Rightarrow$ use $\boldsymbol{T} = \text{diag}(\boldsymbol{Q})$

Stokes preconditioner: $P = \begin{bmatrix} A_{MG} & 0 \\ 0 & \text{diag}(Q) \end{bmatrix}$

Theory: $\lambda(P^{-1}\mathcal{A}) \in [-a, -b] \cup [c, d]$ with a, b, c, d independent of h (*Silvester & W*)

Number of MINRES iterations for 10^{-6} residual reduction (CPU time)
A_{MG}: 1 V-cycle for Laplacian with 1-1 damped Jacobi smoothing

Driven Cavity flow

Grid	Mixed Element				
	Q_1–P_0	Q_2–Q_1	Q_2–P_1	Q_2–P_0	direct
16×16	36 (6)	45 (5)	29 (5)	25 (5)	(.3)
32×32	38 (8)	50 (9)	31 (7)	25 (6)	(3)
64×64	38 (21)	50 (27)	31 (19)	27 (16)	(31)
128×128	37 (76)	49 (102)	29 (69)	27 (59)	(221)
256×256	36 (313)	47 (427)	29 (305)	27 (267)	(8961)

Nonsymmetric definite problems:

in context of (linear) PDEs arise from 1st order derivatives.
main example: Convection-Diffusion:

$$-\nu\nabla^2 u + b.\nabla u = f$$

$\Rightarrow$

$$F\mathbf{u} \equiv (\nu A + N)\mathbf{u} = \mathbf{f}$$

A discrete Laplacian as before, N convection

For Galerkin approximation (central finite differences) A is SPD and N is skew-symmetric if $\nabla \cdot b = 0$ and for certain boundary conditions
$\Rightarrow$

$$\mathbf{u}^T F\mathbf{u} = \nu\mathbf{u}^T A\mathbf{u} \geq 0$$

hence eigenvalues in right half plane.

Nonsymmetric definite problems: Convection-Diffusion:

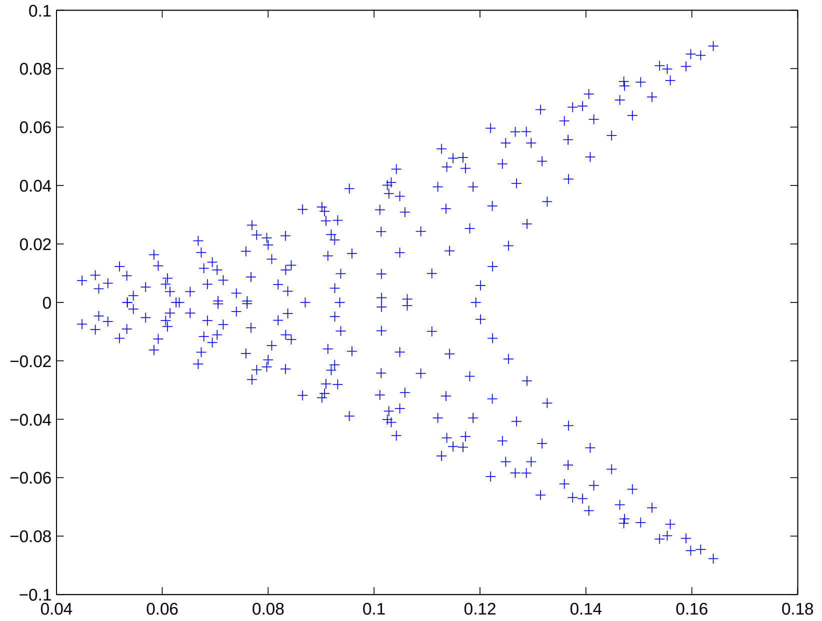
Unfortunately: oscillatory approximations for ν/h small
Key analytic issue: need stable finite element approximation

$\longrightarrow$ SUPG (Streamline upwind Petrov Galerkin)
(*Hughes & Brookes*)

$$\boldsymbol{F}\mathbf{u} \equiv (\boldsymbol{\nu A} + \boldsymbol{N} + \boldsymbol{S})\mathbf{u} = \mathbf{f}$$

$\boldsymbol{S}$: symmetric (semi-definite) stabilisation

eigenvalues:



Nonsymmetric definite problems:

Iterative methods: GMRES , BICGSTAB (ℓ), QMR , CGS , GMRESR ,. . .

Convergence: Let $F = H + S$ where
$H = (F + F^T)/2, S = (F - F^T)/2$ (symmetric—skew)

If H is SPD, GMRES residuals satisfy

$$\frac{\|\mathbf{r}^{(k)}\|}{\|\mathbf{r}^{(0)}\|} \le \left(1 - \frac{\lambda_{\min}(H)^2}{\lambda_{\min}(H)\lambda_{\max}(H) + \rho(S)^2}\right)^{k/2}$$

$\rho(S) =$ eigenvalue of S with maximum modulus.

Nonsymmetric definite problems: Convection-Diffusion:

$(\nu A + N + S)\mathbf{u} = \mathbf{f}$:

- if $\nu \sim \mathcal{O}(1)$ similar to SPD
- if ν small $\Rightarrow$ problem is hard!

Successful solvers tend to rely on ordering discrete variables ‘with the flow’ (ie. along b)

Gauss-Seidel almost as good as anything more complicated with such an ordering!

$\Rightarrow$ multigrid with directional smoothing

Theory: very limited

Nonsymmetric indefinite problems:

Same nonsymmetric iterative methods: some of same issues as symmetric indefinite and some same as nonsymmetric definite

Key example: Incompressible flow: Navier-Stokes and its linearizations

For most other (general) problems heuristic/AMG/ILU/... preconditioning

Steady Incompressible Navier-Stokes:

$$\begin{aligned} -\nu\nabla^2 u + u.\nabla u + \nabla p &= f \\ \nabla \cdot u &= 0 \end{aligned}$$

Mixed finite element approximation:

$$u \approx u_h \in X_h \subset \mathcal{H}^1, \quad p \approx p_h \in M_h \subset L_2$$

(Typical) error estimate: (*Girault & Raviart*)

$$\begin{aligned} \|\nabla(u - u_h)\| &+ \|p - p_h\| \\ &\leq C(\inf_{v_h \in X_h} \|\nabla(u - v_h)\| + \inf_{q_h \in M_h} \|p - q_h\|) \\ &\leq C\, h^2 \left(\|D^3 u\| + \|D^2 p\|\right), \end{aligned}$$

$\Rightarrow$ prefered iterative convergence in

$$\|\mathbf{u} - \mathbf{u}^{(k)}\|_A + \|\mathbf{p} - \mathbf{p}^{(k)}\|_Q$$

Navier-Stokes: discretise and linearise (Oseen or Newton):

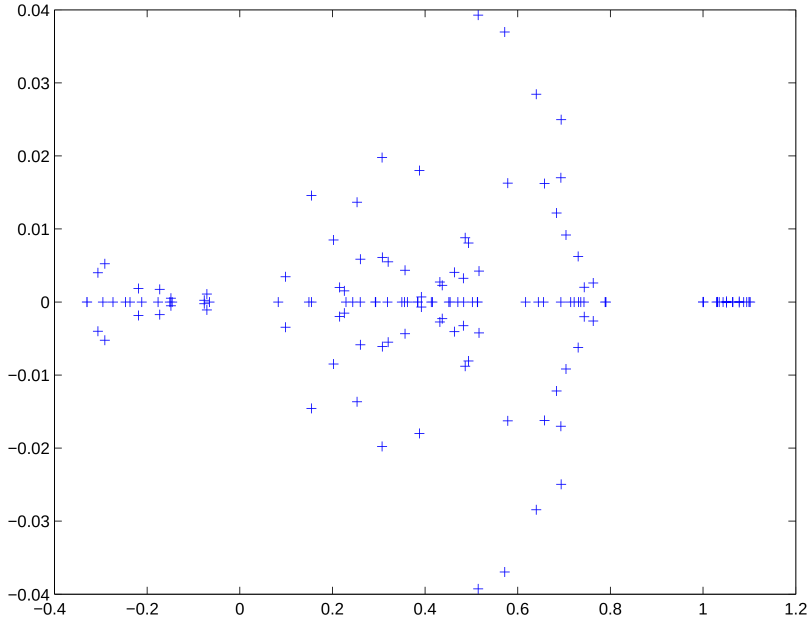
$$\begin{bmatrix} F & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ 0 \end{bmatrix}$$

$\mathbf{u}$: velocity coefficients, $\mathbf{p}$: pressure coefficients

B/B^T: discrete divergence/gradient

$F = \nu A + N (+S)$: discrete convection-diffusion operator
A: discrete (vector) Laplacian, N: convection

eigenvalues:



Oseen/Newton:

$$\mathcal{A} = \begin{bmatrix} F & B^T \\ B & 0 \end{bmatrix}, \quad P = \begin{bmatrix} F_{MG} & B^T \\ 0 & S \end{bmatrix}$$

non-symmetric $\Rightarrow$ use GMRES , BICGSTAB (ℓ), QMR , ...

Observation: the choice $S = BF^{-1}B^T$, $F_{MG} = F$ will give 2 distinct eigenvalues for preconditioned system $\Rightarrow$ convergence in 2 iterations (*Murphy, Golub & W*)

$\Rightarrow$ Schur complement approximations $\rightarrow S$

Preconditioners comprise

- multigrid cycle for $F = \nu A + N$, convection-diffusion
- Schur complement approximation $S \sim BF^{-1}B^T$
- simple multiply by B^T

Schur complement approximations:

Note: $BB^T \sim \nabla \cdot \nabla \sim QA_p$: discrete Laplacian on pressure space

If F_p is similarly a convection-diffusion operator on the pressure space, can expect

$$FB^T \sim B^T F_p$$

$$\Rightarrow BB^T \sim BF^{-1}B^T F_p = SF_p$$

$$\Rightarrow S^{-1} \sim F_p(BB^T)^{-1} \sim F_p A_p^{-1} Q^{-1}$$

Outcome: choose

$$S^{-1} = F_p A_p^{-1} Q^{-1} \qquad \text{(\textit{Kay \& Loghin})}$$

with $A_p^{-1} \longrightarrow$ Laplacian MG cycle and $Q^{-1} \longrightarrow$ diag(Q)

Oseen/Newton:

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \boldsymbol{F} & \boldsymbol{B}^T \\ \boldsymbol{B} & \mathbf{0} \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} \boldsymbol{F}_{MG} & \boldsymbol{B}^T \\ \mathbf{0} & \boldsymbol{S} \end{bmatrix}$$

with

$$\boldsymbol{S}^{-1} = \boldsymbol{F}_p \widehat{\boldsymbol{A}_p^{-1}} \widehat{\boldsymbol{Q}^{-1}}$$

and $\widehat{\boldsymbol{A}_p^{-1}} = 1$ Laplacian MG cycle, $\widehat{\boldsymbol{Q}^{-1}} = \mathbf{1}/\text{diag}(\boldsymbol{Q})$ is the pressure convecton-diffusion preconditioner (FP)

Oseen/Newton:

$$\mathcal{A} = \begin{bmatrix} F & B^T \\ B & 0 \end{bmatrix}, \quad P = \begin{bmatrix} F_{MG} & B^T \\ 0 & S \end{bmatrix}$$

with

$$S^{-1} = F_p \widehat{A_p^{-1}} \widehat{Q^{-1}}$$

and $\widehat{A_p^{-1}} = 1$ Laplacian MG cycle, $\widehat{Q^{-1}} = 1/\mathrm{diag}(Q)$ is the pressure convecton-diffusion preconditioner (FP)

A second approach: (*Elman*)

$$S^{-1} = \widehat{A_p^{-1}} \underbrace{(B\widehat{Q^{-1}}F\widehat{Q^{-1}}B^T)}_{\text{just multiply}} \widehat{A_p^{-1}}$$

gives the least squares commutator preconditioner (BFBT)

Number of GMRES iterations for 10^{-6} residual reduction (CPU time)

pressure convection-diffusion preconditioning with Laplacian algebraic multigrid (`femlab`)

Q_2–Q_1 mixed approximation Oseen Driven Cavity problem from a Picard iterati on

$\mathcal{R} = 2/\nu$ is the Reynolds number

grid	$\mathcal{R}$=10	100	200
8×8	23 (0.4)	29 (0.6)	58 (1.2)
16×16	24 (1.5)	31 (1.9)	41 (2.8)
32×32	23 (5.7)	31 (7.9)	34 (9.2)
64×64	21 (21.6)	28 (29.6)	31 (35.3)
128×128	21 (86.8)	27 (113.6)	31 (168.0)

Number of GMRES iterations for 10^{-6} residual reduction (CPU time)
pressure convection-diffusion preconditioning with Laplacian algebraic multigrid (`femlab`)
Q_2–Q_1 mixed approximation Oseen Driven Cavity problem from a Picard iterati on
$\mathcal{R} = 2/\nu$ is the Reynolds number

grid	$\mathcal{R}$=10	100	200
8×8	23 (0.4)	29 (0.6)	58 (1.2)
16×16	24 (1.5)	31 (1.9)	41 (2.8)
32×32	23 (5.7)	31 (7.9)	34 (9.2)
64×64	21 (21.6)	28 (29.6)	31 (35.3)
128×128	21 (86.8)	27 (113.6)	31 (168.0)

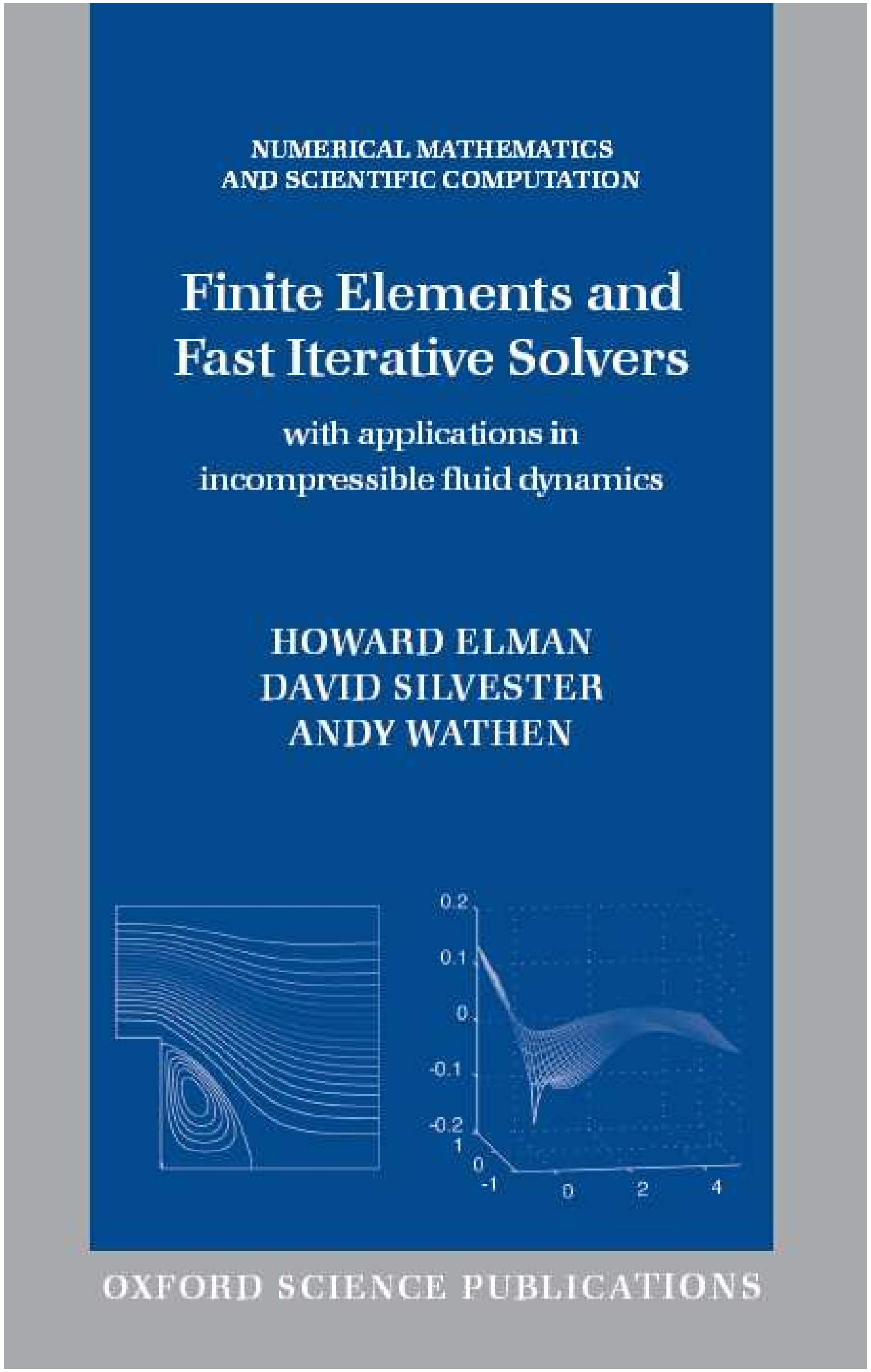
Convergence in right norms?

Have considered:

- Poisson - symmetric positive definite
- Stokes - symmetric indefinite
- Convection-diffusion - nonsymmetric but positive definite
- Navier-Stokes (Oseen/Newton) - nonsymmetric indefinite

and

- constructed optimal preconditioned Krylov subspace iterative solvers through use of multigrid for Laplace and convection-diffusion subproblems
- shown convergence in the relevant norms for the underlying PDE problems


NUMERICAL MATHEMATICS
AND SCIENTIFIC COMPUTATION
Finite Elements and
Fast Iterative Solvers
with applications in
incompressible fluid dynamics
HOWARD ELMAN
DAVID SILVESTER
ANDY WATHEN
0.2
0.1
0
-0.1
-0.2
1
0
-1
0
2
4
OXFORD SCIENCE PUBLICATIONS

`matlab` software: IFISS freely downloadable from

`http://www.manchester.ac.uk/ifiss`
or
`http://www.cs.umd.edu/~elman/ifiss.html`