Subspace Correction Preconditioners

Sébastien Loisel

Heriot-Watt University

May 29, 2012

Outline

Linear problems

$$Ax = b$$

$$\begin{bmatrix} 3 & 2 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 18 \\ 22 \end{bmatrix}$$

Solve using Gaussian Elimination: $E2 := E2 - (1/3)E1$:

$$\begin{bmatrix} 3 & 2 \\ 0 & 5.3333 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 18 \\ 16 \end{bmatrix}$$

Backward substitution:

$$x_2 = 16/5.3333 = 3.0000$$

$$3x + 2 \cdot 3.0000 = 18 \implies x_1 = 4.$$

Solution: $x = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

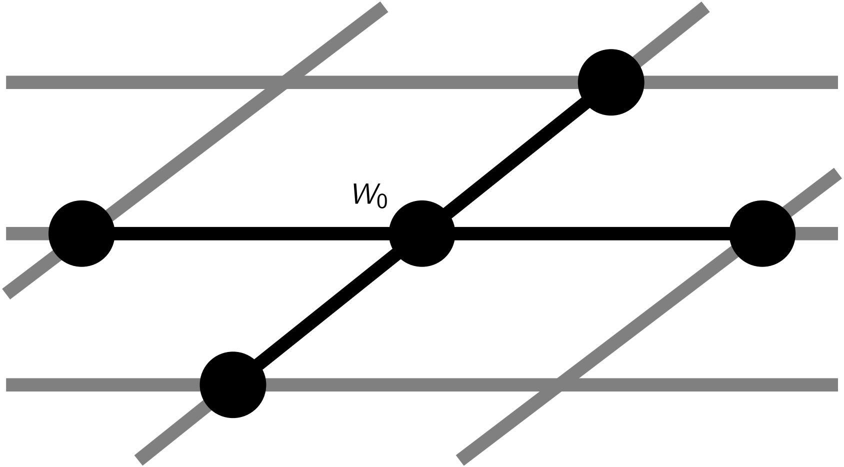
Sparse matrices in MATLAB

```
>> I = eye(3)
I =
  1   0   0
  0   1   0
  0   0   1
>> sparse(I)
ans =
   (1,1)  1
   (2,2)  1
   (3,3)  1
>> I = eye(100000000,100000000);
??? Error using ==> eye
Maximum variable size allowed by the program is exceeded.
>> I = speye(100000000,100000000);
>> trace(I)
ans =
   100000000
```

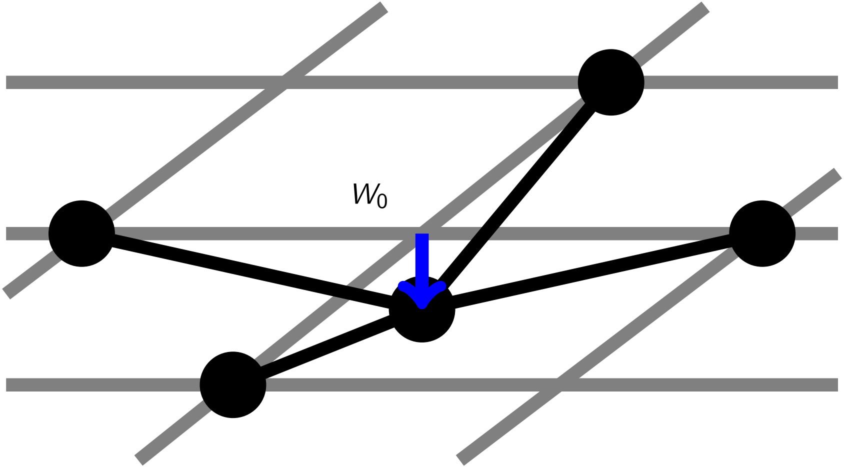
Sparse matrices in MATLAB

```
>> full(delsq(numgrid('L',8)))
ans =
  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
 -1  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0 -1  4  0  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
 -1  0  0  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0 -1  0 -1  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0 -1  0 -1  4  0  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0 -1  0  0  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0 -1  0 -1  4 -1  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0 -1  0 -1  4  0  0 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0 -1  0  0  4 -1  0  0  0  0 -1  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0 -1  0 -1  4 -1  0  0  0  0 -1  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0 -1  0 -1  4 -1  0  0  0  0 -1  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0 -1  4 -1  0  0  0  0 -1  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0 -1  4 -1  0  0  0  0 -1  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  4  0  0  0  0  0 -1  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0 -1  0  0  0  0  0  4 -1  0  0  0  0 -1  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0  0  0 -1  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0  0  0 -1  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0  0  0 -1  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0  0  0 -1  0
  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4  0  0  0  0  0 -1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0  0  4 -1  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1  0
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4 -1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 -1  0  0  0  0 -1  4
```

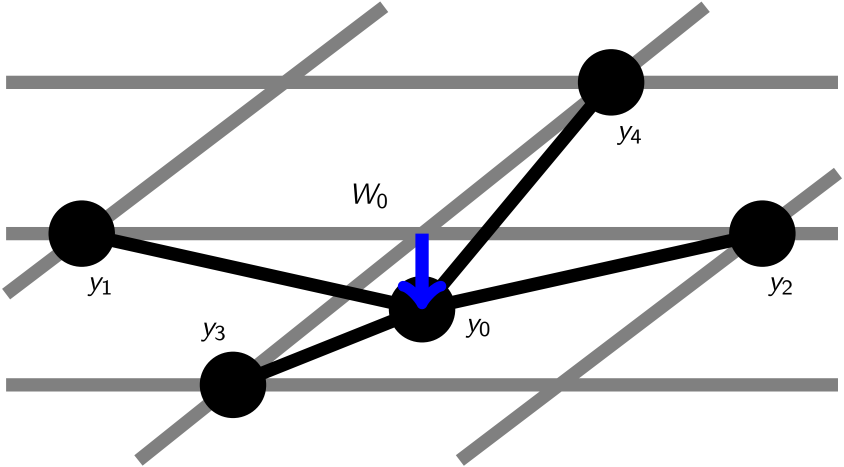
The meaning of the delsq matrix



The meaning of the delsq matrix



The meaning of the delsq matrix

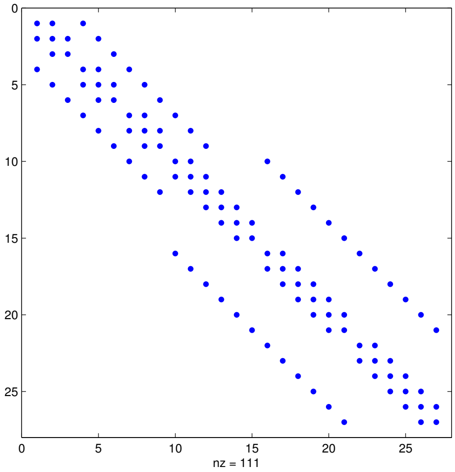


$$\frac{y_1+y_2+y_3+y_4}{4} - y_0 = cW_0$$

`delsq`: the resulting linear system

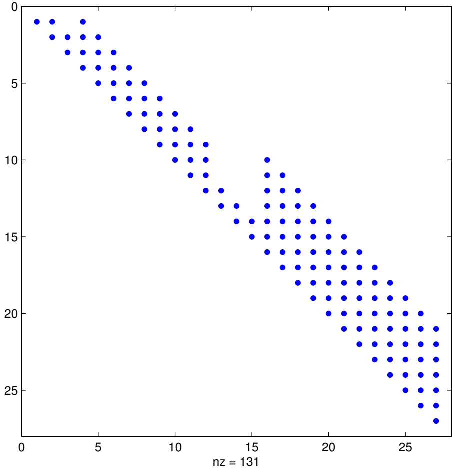
The spy command reveals the sparsity

```
>> spy(delsq(numgrid('L',8)));
```



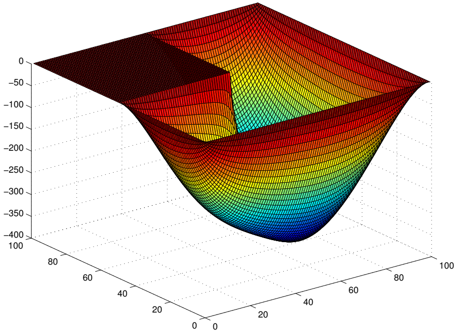
The spy command reveals the fill-in

```
>> [L,U] = spy(lu(delsq(numgrid('L',8))));
>> spy(U);
```



GE on a large enough matrix will exhaust memory

```
>> G = numgrid('L',100);
>> A = delsq(G);
>> u = A \ ones(7203,1);
>> u0 = zeros(size(G));
>> u0(G>0) = u;
>> surf(-u0);
>> nnz(A)
ans =
      35623
>> nnz(lu(A))
ans =
    1164679
```



2. Iterative linear solvers

Geometric series to avoid running out of memory

$$(I - E)^{-1} = \sum_{k=0}^{\infty} E^k$$

Use $E = I - A$ to get

$$A^{-1} = \sum_{k=0}^{\infty} (I - A)^k$$

This will work iff $\rho(E) < 1$ ie.

$$\rho(I - A) < 1.$$

We need $A \approx I$!

```
>> max(abs(eigs(A-speye(size(A)))))
ans =
6.9961
```

Not going to work.

Method of search directions

Idea: pick search directions $v_1, \ldots, v_k$.

$$V = \begin{bmatrix} v_1 & \ldots & v_k \end{bmatrix}.$$

Set $x_k = x_0 + \sum_{i=1}^{k} y_i v_i$ with $y_1, \ldots, y_k$ to be chosen later.

$$x_k = x_0 + Vy.$$

Compute the residual $r_k = b - Ax_k$:

$$r_k = r_0 - AVy.$$

Choose $y_1, \ldots, y_n$ to minimize the residual:

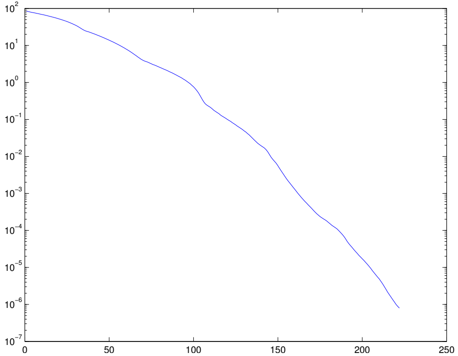
$$\|r_k\|_2 = \min \implies y = ((AV)^T(AV))^{-1}(AV)^T r_0.$$

(Normal Equations; could use other method to solve LS problem.)
How to pick $v_1, \ldots, v_k$?

Krylov search directions

Let $v_0 = r_0$ and $v_{k+1} = Av_k$.
To get a better basis, use Gram-Schmidt on $v_0, \ldots, v_k$.

```
>> [x,f,r,i,rv] = gmres(A,ones(7203,1),[],1e-8,1000);
>> semilogy(rv)
```



GMRES as matrix polynomials

$$v_1 = r_0, \qquad v_k = Av_{k-1} = A^2 v_{k-2} = \ldots = A^k r_0$$

$$r_k = b - Ax_k = b - A(x_0 + Vy) = r_0 - AVy$$

$$= r_0 - \sum_{i=0}^{k-1} p_{i+1} A^{i+1} r_0 = \overbrace{\left(I - \sum_{i=0}^{k-1} p_{i+1} A^{i+1} \right)}^{p(A)} r_0$$

where $p(A)$ is a polynomial of degree k such that $p(0) = 1$ **and such that** $\|r_k\|_2$ **is minimized.** Hence,

$$\|r_k\|_2 \leqslant \||p(A)\||_2 \|r_0\|_2$$

Convergence analysis of GMRES

Theorem

Let A be normal with eigenvalues $\{\lambda_i\}$

$$\|r_k\|_2 \leqslant \|r_0\|_2 \min_{\substack{\deg p = k \\ p(0)=1}} \max_{\lambda \in \{\lambda_i\}} |p(\lambda)|.$$

Proof.

A normal so $A = UDU^T$ where $U^T U = I$ and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

$$A^k = \overbrace{(UDU^T)(UDU^T)\dots(UDU^T)}^{k \text{ terms}} = UD^kU^T$$

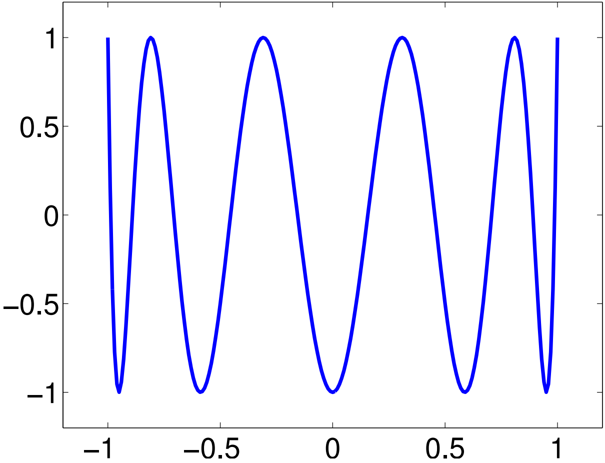
$$p(A) = \sum p_i A^i = U\left(\sum (p_i D^i)\right)U^T = Up(D)U^T$$

$$\|r_k\|_2 \leqslant \|p(A)\|_2 \|r_0\|_2 = \|p(D)\|_2 \|r_0\|_2,$$

as required. □

Chebyshev polynomials

$$T_k(x) = \cosh(k \operatorname{acosh}(x))$$



Use Chebyshev points for spd matrices

Say A is symmetric positive definite with $\{\lambda_i\} \subset [a, b]$.

$$p_k(x) = \frac{T_k\left(2\frac{x-a}{b-a} - 1\right)}{T_k\left(2\frac{-a}{b-a} - 1\right)} \text{ where } T_k(x) = \cosh(k\,\text{acosh}(x))$$

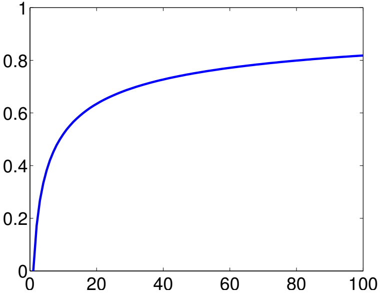
Theorem

Assume A is spd and let $\kappa = b/a$.

$$\|r_k\|_2 \leqslant \|r_0\|_2 \frac{1}{T_k\left(2\frac{-a}{b-a} - 1\right)} \approx \|r_0\|_2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k,$$

Convergence factor

$$\rho(\kappa) = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$$



Small condition number $\kappa \implies$ fast convergence.
Big condition number $\kappa \implies$ slow convergence.

3. Subspace preconditioning

Preconditioning

Let P be a matrix.

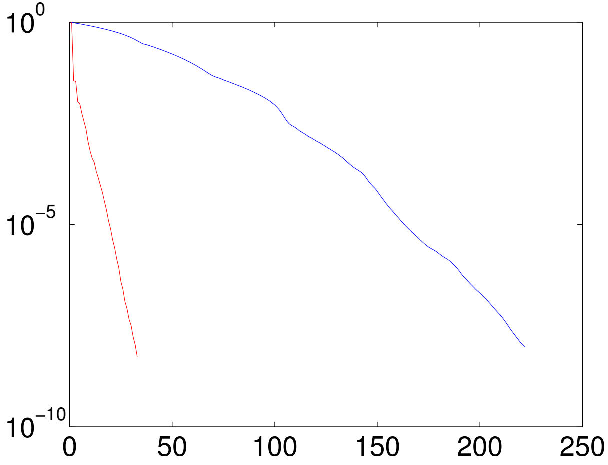
$$Ax = b \text{ is equivalent to } P^{-1}Ax = P^{-1}b.$$

Maybe $\kappa(P^{-1}A) \ll \kappa(A)$.

```
>> P = blkdiag(A(1:3000,1:3000),A(3001:end,3001:end));
>> eigs(P\A,1)/eigs(P\A,1,'sm')
   36.2759
>> eigs(A,1)/eigs(A,1,'sm')
ans =
   2.0668e+03
```

J Xu, Iterative methods by space decomposition and subspace correction. SIREV 34, 1992, pp. 581–613.

Preconditioned GMRES



Residual norms for preconditioned (red) and unpreconditioned (blue) GMRES.

Symmetry and preconditioning

Problem: $P^{-1}A$ is nonsymmetric.

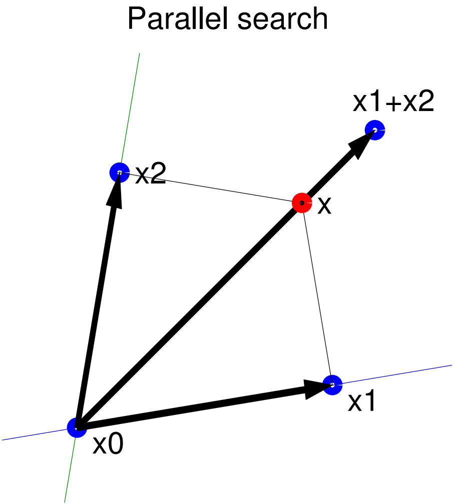
$$p(P^{-1}A)r_0 = A^{-\frac{1}{2}} p\left(A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}\right) \overbrace{A^{\frac{1}{2}} r_0}^{\tilde{r}} .$$

Solution: Minimize the $A-$norm instead of the $2-$norm:

$$\begin{aligned} \|p(P^{-1}A)r_0\|_A^2 &= \left(A^{-\frac{1}{2}} p\left(A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}\right) \tilde{r}\right)^T A \left(A^{-\frac{1}{2}} p\left(A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}\right) \tilde{r}\right) \\ &= \|p\left(A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}\right) \tilde{r}\|_2^2 \end{aligned}$$

and $A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}$ is symmetric. Note that eigenvalues of $A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}$ are the same as for $P^{-1}A$!
Remark: this is how you get PCG from CG.
Remark: Can also use the P-norm – maybe good for GMRES.

Subspace correction



Use A-inner product $\langle u, v \rangle_A := u^T A v$.

Implementing an A-orthogonal projection

Theorem (Projection to the column space of R_0^T)

Let $A_0 = R_0 A R_0^T$. Then $\mathcal{P}_0 = R_0^T A_0^{-1} R_0 A$ is an A-orthonormal projection:

$$\mathcal{P}_0^2 = \mathcal{P}_0 \text{ and } \langle \mathcal{P}_0 u, v \rangle_A = \langle u, \mathcal{P}_0 v \rangle_A$$

Proof.

$$\mathcal{P}_0^2 = R_0^T A_0^{-1} \overbrace{R_0 A R_0^T}^{A_0} A_0^{-1} R_0 A = R_0^T A_0^{-1} \cancel{A_0 A_0^{-1}} R_0 A = \mathcal{P}_0$$
$$\langle u, \mathcal{P}_0 v \rangle_A = u^T A R_0^T A_0^{-1} R_0 A v = \langle \mathcal{P}_0 u, v \rangle_A$$
□

Parallel subspace correction: $P^{-1}A = \sum_{i=1}^m R_i^T A_i^{-1} R_i A = \sum_{i=1}^m \mathcal{P}_i$.

Jacobi example: 7 subspaces

The Jacobi preconditioner is the simplest parallel subspace correction preconditioner.

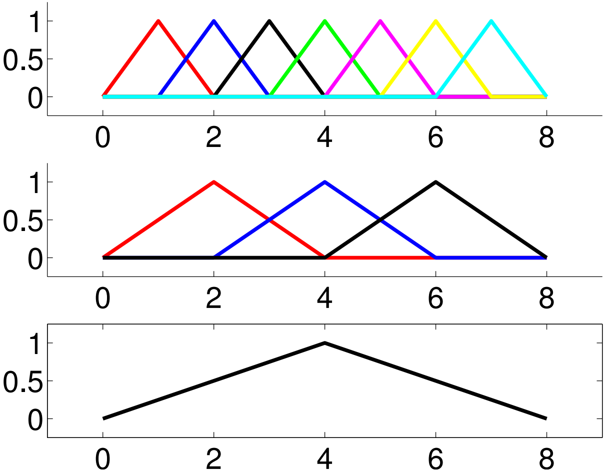
$$R_1^T = \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\end{bmatrix} R_2^T = \begin{bmatrix}0\\1\\0\\0\\0\\0\\0\end{bmatrix} R_3^T = \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\end{bmatrix} R_4^T = \begin{bmatrix}0\\0\\0\\1\\0\\0\\0\end{bmatrix} R_5^T = \begin{bmatrix}0\\0\\0\\0\\1\\0\\0\end{bmatrix} R_6^T = \begin{bmatrix}0\\0\\0\\0\\0\\1\\0\end{bmatrix} R_7^T = \begin{bmatrix}0\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

$$A = \begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{bmatrix}; \quad \begin{matrix} A_1 = A_2 = \ldots = A_7 = 2; \\ P^{-1} = \sum_i R_i A_i^{-1} R_i^T = 0.5I \end{matrix}$$

$$\kappa(A) = 25.27 \qquad \kappa(P^{-1}A) = 25.27.$$

That did not help **at all**.

Multigrid example: 11 subspaces



Each subspace is $V_i = \operatorname{span} \phi_i$ (a 1d subspace of $H_0^1(0, 8)$); $i = 1, 2, \ldots, 11$.

Multigrid example, continued

$$R_1^T = \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\end{bmatrix} R_2^T = \begin{bmatrix}0\\1\\0\\0\\0\\0\\0\end{bmatrix} R_3^T = \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\end{bmatrix} R_4^T = \begin{bmatrix}0\\0\\0\\1\\0\\0\\0\end{bmatrix} R_5^T = \begin{bmatrix}0\\0\\0\\0\\1\\0\\0\end{bmatrix} R_6^T = \begin{bmatrix}0\\0\\0\\0\\0\\1\\0\end{bmatrix} R_7^T = \begin{bmatrix}0\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

$$R_8^T = \begin{bmatrix}0.5\\1\\0.5\\0\\0\\0\\0\end{bmatrix} R_9^T = \begin{bmatrix}0\\0\\0.5\\1\\0.5\\0\\0\end{bmatrix} R_{10}^T = \begin{bmatrix}0\\0\\0\\0\\0.5\\1\\0.5\end{bmatrix} \qquad R_{11}^T = \begin{bmatrix}0.25\\0.5\\0.75\\1\\0.75\\0.5\\0.25\end{bmatrix}$$

$A = \text{tridiag}(-1, 2, -1)$; $A_1 = A_2 = \ldots = A_7 = 2$; $A_8 = A_9 = A_{10} = 1$; $A_{11} = 0.5$.

$$\kappa(A) = 25.27 \qquad \kappa(P^{-1}A) = 2.866.$$

Domain decomposition

$$R_1^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad R_2^T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A_1 = A_2 = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} 0.8 & 0.6 & 0.4 & 0.2 & & & \\ 0.6 & 1.2 & 0.8 & 0.4 & & & \\ 0.4 & 0.8 & 1.2 & 0.6 & & & \\ 0.2 & 0.4 & 0.6 & 1.6 & 0.6 & 0.4 & 0.2 \\ & & & 0.6 & 1.2 & 0.8 & 0.4 \\ & & & 0.4 & 0.8 & 1.2 & 0.6 \\ & & & 0.2 & 0.4 & 0.6 & 0.8 \end{pmatrix}$$

$$\kappa(A) = 25.27 \qquad \kappa(P^{-1}A) = 7.23.$$

4. Analysis

Definition (Stable decomposition [SD])

Let $K_0 < \infty$. A stable decomposition of v is a choice $\{v_i\}$ such that $v = \sum_{i=1}^m \mathcal{P}_i v_i$ and $\sum_{i=1}^m \|v_i\|_A^2 \leqslant K_0 \|v\|_A^2$. We assume all v have a stable decomposition.

Example

Let $v \in H_0^1(a,b)$ be piecewise linear (grid param h) and let $v_i = v$ in $\Omega_1 = (a,c) \subset (a,b)$ and $v = 0$ outside (a,c).
Trace th: $|v(c-h)| < C|v|_{H_0^1(a,b)}$ hence $|v'(x)| \leqslant \frac{C}{h}|v_1|_{H_0^1(a,b)}$ for $x \in (c-h,c)$. Hence: $\int_a^c |v'(x)|^2 \leqslant (1 + C^2 h^{-1})|v|_{H_0^1(a,b)}^2$.
Hence:

$$K_0 \leqslant 1 + 2C^2 h^{-1}$$

Definition (Limited redundancy [LR])

For all $S \subset \{1,\ldots,m\} \times \{1,\ldots,m\}$:

$$\sum_{(i,j)\in S} \langle \mathcal{P}_i v_i, \mathcal{P}_j w_i \rangle_A \leqslant K_1 \left(\sum_{i=1}^{m} \langle \mathcal{P}_i v_i, v_i \rangle_A \right)^{\frac{1}{2}} \left(\sum_{i=1}^{m} \langle \mathcal{P}_i w_i, w_i \rangle_A \right)^{\frac{1}{2}}.$$

Example

Say $\langle \mathcal{P}_i v_i, \mathcal{P}_j w_j \rangle_A \leqslant \epsilon_{ij} \|\mathcal{P}_i v_i\|_A \|\mathcal{P}_j v_j\|_A$ for all v_i, w_j; then:

$$\sum_{(i,j)\in S} \langle \mathcal{P}_i v_i, \mathcal{P}_j w_j \rangle_A \overset{CS}{\leqslant} \sum_{(i,j)\in S} \epsilon_{ij} \|\mathcal{P}_i v_i\|_A \|\mathcal{P}_j w_j\|_A \leqslant \rho(\epsilon) \left(\sum_{i=1}^{m} \|\mathcal{P}_i v_i\|_A^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{m} \|\mathcal{P}_j w_j\|_A \right)^{\frac{1}{2}}$$

Since $|\epsilon_{ij}| \leqslant 1$ and ϵ is $m \times m$, we have $K_1 \leqslant |||\epsilon|||_\infty \leqslant m$.

- In DD and MG, most $\epsilon_{ij} = 0$ since the subspaces V_i and V_j are "far apart" and thus $\rho(\epsilon) = O(\log m)$.

Theorem (Analysis of parallel subspace correction)

Assume [SD] and [LR]. Then, $\kappa(P^{-1}A) \leqslant K_0 K_1$.

$$[SD] : \sum_{i=1}^{m} \|v_i\|_A^2 \leqslant K_0 \|v\|_A^2 \qquad [LR] : \sum_{ij} \langle \mathcal{P}_i w, \mathcal{P}_j w \rangle_A \leqslant K_1 \sum_i \langle \mathcal{P}_i w, w \rangle_A.$$

Proof.

Upper bound for $\|P^{-1}Av\|_A^2 = \|\sum_{i=1}^n \mathcal{P}_i v\|_A^2$:

$$\|P^{-1}Av\|_A^2 = \langle \sum_i \mathcal{P}_i v, \sum_j \mathcal{P}_j v \rangle_A = \sum_{ij} \langle \mathcal{P}_i v, \mathcal{P}_j v \rangle_A$$

$$\overset{[LR]}{=} K_1 \sum_i \langle \mathcal{P}_i v, v \rangle_A = K_1 \langle P^{-1}Av, v \rangle_A \overset{CS}{\leqslant} K_1 \|P^{-1}Av\|_A \|v\|_A;$$

$\|P^{-1}Av\|_A \leqslant K_1 \|v\|_A$ and hence $\lambda_{\max}(P^{-1}A) \leqslant K_1$.

□

Theorem (Analysis of parallel subspace correction)

Assume [SD] and [LR]. Then, $\kappa(P^{-1}A) \leqslant K_0 K_1$.

$$[SD] : \sum_{i=1}^{m} \|v_i\|_A^2 \leqslant K_0 \|v\|_A^2 \qquad [LR] : \sum_{ij} \langle \mathcal{P}_i w, \mathcal{P}_j w \rangle_A \leqslant K_1 \sum_i \langle \mathcal{P}_i w, w \rangle_A .$$

Proof.

Use stable decomposition to find upper bound for $\|v\|_A^2 = \langle v, v \rangle_A$:

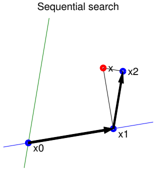
$$\|v\|_A^2 = \sum_i \langle \mathcal{P}_i v_i, v \rangle_A = \sum_i \langle \mathcal{P}_i v_i, \mathcal{P}_i v \rangle_A$$

$$\overset{CS}{\leqslant} \left(\sum_i \langle \mathcal{P}_i v_i, \mathcal{P}_i v_i \rangle_A \right)^{\frac{1}{2}} \left(\sum_i \langle \mathcal{P}_i v, \mathcal{P}_i v \rangle_A \right)^{\frac{1}{2}}$$

$$\overset{[SD]}{\leqslant} \sqrt{K_0} \|v\|_A \left(\langle \sum_i \mathcal{P}_i v, v \rangle_A \right)^{\frac{1}{2}} \overset{CS}{\leqslant} \sqrt{K_0} \|v\|_A \left(\| \sum_i \mathcal{P}_i v \|_A \|v\|_A \right)^{\frac{1}{2}}$$

as required. □

Sequential subspace correction



$$x_{k+\frac{i}{m}} = x_{k+\frac{i-1}{m}} + \mathcal{P}_i(x - x_{k+\frac{i-1}{m}}); \text{ hence:}$$

$$\overbrace{(x - x_{k+\frac{i}{m}})}^{e_{k+\frac{i}{m}}} = \overbrace{(x - x_{k+\frac{i-1}{m}})}^{e_{k+\frac{i-1}{m}}} - \mathcal{P}_i \overbrace{(x - x_{k+\frac{i-1}{m}})}^{e_{k+\frac{i-1}{m}}} = (I - \mathcal{P}_i)e_{k+\frac{i-1}{m}} \text{ thus}$$

$$e_{k+1} = \overbrace{\left(\prod_{i=1}^{m}(I - \mathcal{P}_i)\right)}^{E_m} e_k$$

E_m is the "error iterator". The convergence factor is $\|E_m\|_A$.

Theorem

$$\|\|E_m\|\|_A \leqslant 1 - \frac{1}{K_0(1+K_1)^2} =: \delta < 1 \textit{ with } E_m = \prod_{i=1}^{m}(I - \mathcal{P}_i).$$

Proof.

Algebra: $I - E_{i-1} = \sum_{j=1}^{i-1} \mathcal{P}_j E_{j-1}$. From the parallel method, we know that $\|v\|_A \leqslant K_0 \| \sum_{i=1}^m \mathcal{P}_i v\|_A$. Because $\sum_i \mathcal{P}_i$ is A-symmetric, this gives the Rayleigh quotient bound $\langle \sum_i \mathcal{P}_i v, v\rangle_A / \langle v, v\rangle_A \leqslant K_0$.

$$\sum_{i=1}^m \langle \mathcal{P}_i v, v\rangle_A = \overbrace{\sum_{i=1}^m \langle \mathcal{P}_i v, E_{i-1} v\rangle_A}^{X} + \overbrace{\sum_{i=1}^m \langle \mathcal{P}_i v, (I - E_{i-1}) v\rangle_A}^{Y = \sum_{i=1}^m \langle \mathcal{P}_i v, \sum_{j=1}^{i-1} \mathcal{P}_j E_{j-1} v\rangle_A}$$

$$X \overset{CS}{\leqslant} \left(\sum_{i=1}^m \langle \mathcal{P}_i v, v\rangle_A\right)^{\frac{1}{2}} \left(\sum_{i=1}^m \langle \mathcal{P}_i E_{i-1} v, E_{i-1} v\rangle_A\right)^{\frac{1}{2}}$$

$$Y \overset{[LR]}{\leqslant} K_1 \left(\sum_{i=1}^m \langle \mathcal{P}_i v, v\rangle_A\right)^{\frac{1}{2}} \left(\sum_{i=1}^m \langle \mathcal{P}_i E_{i-1} v, E_{i-1} v\rangle_A\right)^{\frac{1}{2}}$$

□

Theorem

$$\|\|E_m\|\|_A \leqslant 1 - \frac{1}{K_0(1+K_1)^2} =: \delta < 1 \text{ with } E_m = \prod_{i=1}^{m}(I - \mathcal{P}_i).$$

Proof.

$$\sum_{i=1}^{m} \langle \mathcal{P}_i E_{i-1} v, E_{i-1} v \rangle_A = \|v\|_A^2 - \|E_m v\|_A^2$$

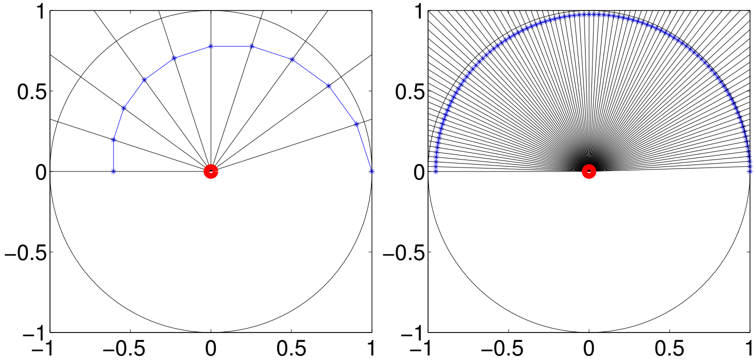
Check:

$$\begin{aligned} \|E_{i-1}v\|_A^2 - \|E_i v\|_A^2 &= \langle E_{i-1}v, E_{i-1}v\rangle_A - \langle (I-\mathcal{P}_i)E_{i-1}v, (I-\mathcal{P}_i)E_{i-1}v\rangle_A \\ &= \langle \mathcal{P}_i E_{i-1}v, E_{i-1}v\rangle_A. \qquad \text{Hence:} \\ \|v\|_A^2 \leqslant K_0(1+K_1)^2(\|v\|_A^2 - \|E_m v\|_A^2). \end{aligned}$$

Rearranging terms, we are done. □

$$\mathcal{P}_i = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix} \text{ where } c = \cos\theta_i \text{ and } s = \sin\theta_i.$$

Take $\theta_i = \pi \frac{i}{m}$. If $\theta_i - \theta_j = \pi/2$ then $\mathcal{P}_i + \mathcal{P}_j = I$ but:



$$\kappa_{par} = 1 \text{ while } |||E_m||| \to 1.$$

Dirichlet principle vs Maximum principle

- We have used the energy norm $\|\cdot\|_A$ – implicitly, the Dirichlet principle.
- We could have used instead the maximum principle. This gives similar preconditioners but different analysis.
- Relevant theory: M-matrix theory.

Conclusions

- Solve large linear problems iteratively.
- Preconditioning helps GMRES.
- Standard preconditioners are subspace correction methods.
- Convergence analysis based on sequential or parallel energy norm minimization.