Problem

Consider a matrix problem (s.p.d.) of the form

$$A\boldsymbol{u} = \boldsymbol{f}, \quad A \in \mathbb{R}^{n \times n}$$

Suppose we have a multilevel iteration process $\mathcal{M}_{ ext{MG}}$

$$I - \mathcal{M}_{\mathrm{MG}} A = \left(I - M^{ op} A
ight)^{
u_{\mathrm{pre}}} \left(I - P ig(P^{ op} A Pig)^{-1} P^{ op} Aig) (I - M A)^{
u_{\mathrm{post}}}$$

so that

$$oldsymbol{e} \leftarrow (I - \mathcal{M}_{\mathrm{MG}} A) oldsymbol{e}$$

The iteration converges for any $m{f}$ and $m{u}_0$ iff $ho(I-\mathcal{M}_{\mathrm{MG}}A)<1.$

we are seeking a method that yields **a bound on the error reduction in each iteration that is independent** of n.

Notation

Fine grid $\Omega=\{1,\ldots,n\}=C\cup F$ and coarse grid $\Omega_c=C$. Interpolation / restriction:

$$P:\Omega_c o \Omega \quad ext{ and } \quad R:\Omega o \Omega_c$$

A is s.p.d., $D = \operatorname{diag}(A)$ — defining an inner product:

$$egin{aligned} \langle oldsymbol{u}, oldsymbol{v}
angle_A &= \langle Aoldsymbol{u}, oldsymbol{v}
angle \ \langle oldsymbol{u}, oldsymbol{v}
angle_D &= \langle Doldsymbol{u}, oldsymbol{v}
angle \ \langle oldsymbol{u}, oldsymbol{v}
angle_{AD^{-1}A} &= \langle D^{-1}Aoldsymbol{u}, Aoldsymbol{v}
angle \end{aligned}$$

Define S as post-relaxation

(and \mathcal{S} as the affine version $oldsymbol{u} \leftarrow \mathcal{S}(oldsymbol{u}, oldsymbol{f})$):

$$oldsymbol{u} \leftarrow Soldsymbol{u} + (I-S)A^{-1}oldsymbol{f} \quad ext{or} \quad oldsymbol{e} \leftarrow Soldsymbol{e}$$

Algorithm

$$egin{aligned} oldsymbol{u} \leftarrow \hat{\mathcal{S}}(oldsymbol{u}, oldsymbol{f}) \ oldsymbol{r}_c \leftarrow Roldsymbol{r} \ oldsymbol{e}_c \leftarrow A_c^{-1}oldsymbol{r}_c \ \hat{oldsymbol{e}} \leftarrow Poldsymbol{e}_c \ oldsymbol{u} \leftarrow U + \hat{oldsymbol{e}} \ oldsymbol{u} \leftarrow \mathcal{S}(oldsymbol{u}, oldsymbol{f}) \end{aligned}$$

In general, consider relaxation as

$$S = I - MA$$

Assumptions:

- ullet M is norm convergent (in A): $\|S\|_A < 1$
- \bullet P is full rank

ullet A is s.p.d.

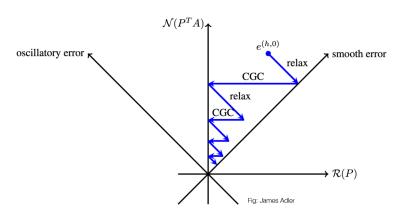
Let the coarse grid correction step be

$$T = I - P(P^{\top}AP)^{-1}P^{\top}A$$

T is an A-orthogonal projection onto the range of P: After coarse grid correction, the error is minimized in the energy norm over $\mathrm{range}(P)$

if
$$w$$
 s.t. $Aw \in \mathcal{N}\left(P^{T}\right)$
$$\left(I - P(P^{T}AP)^{-1}P^{T}A\right)w = w$$

$$\|\cdot\| \ge 1$$



Focus on V(0,1)

The A-adjoint of ST is

$$TS^+ \quad S^+ = I - M^ op A$$

The symmetric V(1,1) cycle is

$$(I - MA) \left(I - P(P^TAP)^{-1}P^TA\right) \left(I - M^TA\right) = STS^+$$

= $STTS^+$

Since $\|ST\|_A = \|TS^+\|_A$ (A-adjoints) we have

$$\left\|STS^{+}
ight\|_{A}=\left\|ST
ight\|_{A}^{2}$$

Ok, so we can focus focus on the V(0,1) cycle, the other cycles follow.

Upper Bound

We will measure convergence or reduction in $\|e\|_A$. Note:

$$\|m{e}\|_A^2 = \|(I-T)m{e}\|_A^2 + \|Tm{e}\|_A^2$$

For a V(0,1) cycle, the reduction in $oldsymbol{e}$ is

$$||STe||_{\Delta}^{2} \leq (1 - \delta^{*})||e||_{\Delta}^{2}$$

we seek a sharp bound

$$\|ST\|_A^2 := \sup_{e
eq 0} rac{\|STe\|_A^2}{\|e\|_A^2} = 1 - \delta^*$$

Sufficient conditions

What should we assume on relaxation and interpolation?

• relaxation is effective on the range of coarse grid correction

There exists $\delta > 0$ such that

$$||STe||_A^2 \leq (1-\delta)||Te||_A^2$$
 for all e .

Then, since T is an A-orthogonal projector

$$\|SToldsymbol{e}\|_A^2 \leq (1-\delta)\|oldsymbol{e}\|_A^2 ext{ for all } oldsymbol{e}$$

• No side effects on the range of interpolation

$$||S\boldsymbol{v}||_A^2 \leq ||\boldsymbol{v}||_A^2$$
 for all $\boldsymbol{v} \perp \operatorname{range}(T)$

Combine these into an assumption

$$||Sv||_A^2 \le ||STv||_A^2 + ||S(I-T)v||_A^2 \le (1-\delta)||Tv||_A^2 + ||(I-T)v||_A^2 = ||v||_A^2 - \delta||Tv||_A^2$$

Theorem

If there exists $\delta > 0$ so that

$$||Se||_A^2 \le ||e||_A^2 - \delta ||Te||_A^2$$
 for all e ,

then

$$\|ST\|_A^2 \le 1 - \delta$$

To be sharp, the largest δ , say $\hat{\delta}$

$$\hat{\delta} = \inf_{e:Te
eq 0} rac{\|e\|_A^2 - \|Se\|_A^2}{\|Te\|_A^2},$$

should be δ^* .

Proof

Since $Toldsymbol{e}=oldsymbol{0}$ gives $\|SToldsymbol{e}\|_A=0$:

$$\|ST\|_A^2 = \sup_{e:Te
eq 0} rac{\|STe\|_A^2}{\|e\|_A^2} = \sup_{e:Te
eq 0} rac{\|STe\|_A^2}{\|Te\|_A^2 + \|(I-T)e\|_A^2}$$

Let \hat{e} be the $\arg\sup$

Then $T\hat{e}$ is also at the supremum.

Thus we have an error at the supremum with $(I-T)\hat{e}=0$

$$\|ST\|_A^2 = \sup_{e:Te \neq 0} \frac{\|STe\|_A^2}{\|e\|_A^2} = \sup_{e:Te \neq 0} \frac{\|S(Te + (I-T)e)\|_A^2}{\|Te\|_A^2} = \sup_{e:Te \neq 0} \frac{\|Se\|_A^2}{\|Te\|_A^2},$$

And

$$1 - \|ST\|_A^2 = \inf_{e: Te \neq \mathbf{0}} \frac{\|Te\|_A^2 - \|Se\|_A^2}{\|Te\|_A^2} = \inf_{e: Te \neq \mathbf{0}} \frac{\|e\|_A^2 - \|Se\|_A^2}{\|Te\|_A^2} = \hat{\delta}$$

The worst δ is sharp

$$\hat{\delta} = \inf_{e: Te
eq 0} rac{\|e\|_A^2 - \|Se\|_A^2}{\|Te\|_A^2},$$

The early theory split this in two part

For some $g(\boldsymbol{e})$ define δ, α_g , and β_g as in

$$\delta(oldsymbol{e}) = \underbrace{rac{\|oldsymbol{e}\|_A^2 - \|Soldsymbol{e}\|_A^2}{g(oldsymbol{e})}}_{lpha_g(oldsymbol{e})} \underbrace{rac{\|oldsymbol{e}\|_A^2}{\|Toldsymbol{e}\|_A^2}}_{1/eta_g(oldsymbol{e})}$$

Consider the smallest α_q and the largest β_g :

$$\hat{lpha}_g = \inf_{oldsymbol{e}: g(oldsymbol{e})
eq 0} lpha_g(oldsymbol{e}) \quad \hat{eta}_g = \sup_{oldsymbol{e}: g(oldsymbol{e})
eq 0} eta_g(oldsymbol{e})$$

For e such that $g(Te) \neq 0$,

$$egin{aligned} \|SToldsymbol{e}\|_A^2 &\leq \|Toldsymbol{e}\|_A^2 - \hat{lpha}_g g(Toldsymbol{e}) \leq \|Toldsymbol{e}\|_A^2 - rac{\hat{lpha}_g}{\hat{eta}_g} \|Toldsymbol{e}\|_A^2 &= \left(1 - rac{\hat{lpha}_g}{\hat{eta}_g}
ight) \|Toldsymbol{e}\|_A^2 \ &\leq \left(1 - rac{\hat{lpha}_g}{\hat{eta}_g}
ight) \|oldsymbol{e}\|_A^2 \end{aligned}$$

Ok, so this is generally worse than the sharp bound

$$\|ST\|_A = \sqrt{1 - \hat{\delta}} \leq \sqrt{1 - rac{\hat{lpha}_g}{\hat{eta}_g}}$$

($lpha_g$ and eta_g are not generally simultaneously satisfied)

Early works, e.g. Ruge-Stüben 1987, use $g(m{e}) = \|m{e}\|_{AD^{-1}A}^2$ Or the weaker form $g(m{e}) = \|Tm{e}\|_{AD^{-1}A}^2$

Smoothing and Approximation

if there exists $ar{lpha}_g>0$ such that

$$||Se||_A^2 \le ||e||_A^2 - \bar{\alpha}_g g(e)$$
 for all e (smoothing)

and there exists $ar{eta}_g>0$ such that

$$\|Toldsymbol{e}\|_A^2 \leq ar{eta}_g g(Toldsymbol{e}) \quad ext{ for all } oldsymbol{e} \quad ext{ (approximation)}$$

then
$$\|ST\|_A \leq \sqrt{1-ar{lpha}_g/ar{eta}_g}$$

Select
$$g(oldsymbol{e}) = \|oldsymbol{e}\|_{AD^{-1}A}^2$$

Since T is an A-orthogonal projection we have

$$\|Te\|_A = \inf_{e_c} \|e - Pe_c\|_A$$

(strong approximation) Assume there is a $ar{eta}_s$ such that

$$\inf_{oldsymbol{e}_c} \|oldsymbol{e} - Poldsymbol{e}_c\|_A^2 \leq ar{eta}_s \|oldsymbol{e}\|_{AD^{-1}A}^2 \quad ext{ for all } oldsymbol{e}.$$

The weaker version looks like (for some \hat{eta})

$$\|Te\|_A^2 \leq \bar{\beta} \|Te\|_{AD^{-1}A} \quad ext{ for all } e.$$

Weaker, means weaker norm. And we can make this a bit more practical. The range of T is A-orthogonal to the range of P, so

$$egin{aligned} \|Toldsymbol{e}\|_A^2 &= \langle AToldsymbol{e}, Toldsymbol{e}
angle &= \langle AToldsymbol{e}, Toldsymbol{e} - Poldsymbol{e}_c
angle \ &\leq \|Toldsymbol{e}\|_{AD^{-1}A}\|Toldsymbol{e} - Poldsymbol{e}_c\|_D. \end{aligned}$$

(weak approximation) Assume that

$$\inf_{oldsymbol{e}_c} \|oldsymbol{e} - Poldsymbol{e}_c\|_D^2 \leq ar{eta}_w \|oldsymbol{e}\|_A^2 \quad ext{ for all } oldsymbol{e},$$