

# Lecture 18: Multigrid



School of Mathematical Sciences, Xiamen University

## 1. Finite difference discretization of a BVP

- Consider the following 1-D Dirichlet boundary value problem

$$\begin{cases} -u''(x) = f(x), & x \in (0, 1) \\ u(0) = u(1) = 0 \end{cases}$$

- For  $n \in \mathbb{N}$ , let

$$h = \frac{1}{n+1}, \quad x_i = ih = \frac{i}{n+1}, \quad 0 \leq i \leq n+1.$$

- The finite difference method is: let  $u_0^h = u_{n+1}^h = 0$ ,  $f_i^h = f(x_i)$ ,  $1 \leq i \leq n$ , find

$$\mathbf{u}^h = [u_1^h \quad u_2^h \quad \cdots \quad u_n^h]^\top$$

such that

$$-\frac{u_{i+1}^h - 2u_i^h + u_{i-1}^h}{h^2} = f_i^h, \quad 1 \leq i \leq n.$$

- The FD system  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$ :

$$\frac{1}{h^2} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1^h \\ u_2^h \\ \vdots \\ u_{n-1}^h \\ u_n^h \end{bmatrix} = \begin{bmatrix} f_1^h \\ f_2^h \\ \vdots \\ f_{n-1}^h \\ f_n^h \end{bmatrix}.$$

## 2. Classical stationary iterative methods (Lecture 6)

- Given a starting vector  $\mathbf{u}^{(0)}$ ,

$$\mathbf{u}^{(j)} = \mathbf{R}\mathbf{u}^{(j-1)} + \mathbf{c}, \quad j = 1, 2, \dots$$

- Jacobi's method
- Gauss–Seidel method
- Successive overrelaxation:  $\text{SOR}(\omega)$
- Symmetric successive overrelaxation:  $\text{SSOR}(\omega)$

## 2.1. Jacobi's method and its relaxation for the FD system

- The iteration matrix of Jacobi's method for  $\mathbf{A}\mathbf{u} = \mathbf{f}$ :

$$\begin{aligned}\mathbf{R} = \mathbf{D}^{-1}(\mathbf{D} - \mathbf{A}) &= \begin{bmatrix} \frac{1}{2} & & & & \\ & \frac{1}{2} & & & \\ & & \ddots & & \\ & & & \frac{1}{2} & \\ & & & & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{2} & 0 & \frac{1}{2} \\ & & & \frac{1}{2} & 0 \end{bmatrix},\end{aligned}$$

$$\mathbf{c} = \mathbf{D}^{-1}\mathbf{f}.$$

- The eigenvalues of  $\mathbf{R}$  are given by

$$\lambda_k = \cos(k\pi h), \quad 1 \leq k \leq n,$$

and the corresponding eigenvectors are given by

$$\mathbf{v}_k = [\sin(k\pi h) \quad \sin(2k\pi h) \quad \cdots \quad \sin(nk\pi h)]^\top, \quad 1 \leq k \leq n.$$

The convergence of the Jacobi's method becomes worse for larger  $n$  since the spectral radius approaches 1 in this situation.

- The relaxation of Jacobi's method ( $\omega\mathbf{A} = \mathbf{D} - (\mathbf{D} - \omega\mathbf{A})$ ):

$$\mathbf{u}^{(j)} = \mathbf{R}(\omega)\mathbf{u}^{(j-1)} + \omega\mathbf{c}, \quad \mathbf{R}(\omega) = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A} = (1 - \omega)\mathbf{I} + \omega\mathbf{R}.$$

The eigenvalues of  $\mathbf{R}(\omega)$  are given by

$$\lambda_k(\omega) = 1 - \omega + \omega\lambda_k = 1 - \omega + \omega\cos(k\pi h), \quad 1 \leq k \leq n.$$

$\mathbf{R}(\omega)$  and  $\mathbf{R}$  have the same eigenvectors.

## 2.2. What makes the convergence of Jacobi's method slow?

- Recall that the solution  $\mathbf{u}$  of the linear system is a fixed point, i.e.,

$$\mathbf{u} = \mathbf{R}\mathbf{u} + \mathbf{c}.$$

This leads to

$$\mathbf{u} - \mathbf{u}^{(j)} = \mathbf{R}(\mathbf{u} - \mathbf{u}^{(j-1)}) = \dots = \mathbf{R}^j(\mathbf{u} - \mathbf{u}^{(0)}).$$

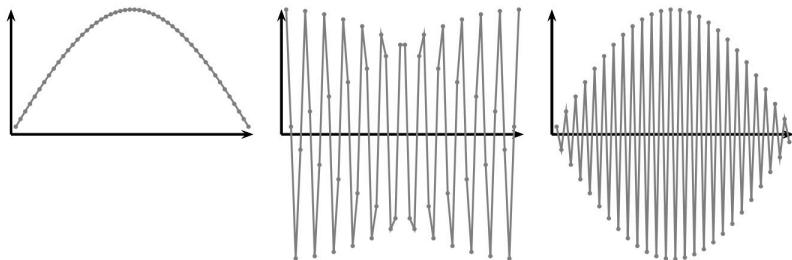
We expand  $\mathbf{u} - \mathbf{u}^{(0)}$  in the basis consisting of the eigenvectors:

$$\mathbf{u} - \mathbf{u}^{(0)} = \sum_{k=1}^n \alpha_k \mathbf{v}_k.$$

This gives

$$\mathbf{u} - \mathbf{u}^{(j)} = \sum_{k=1}^n \alpha_k \lambda_k^j \mathbf{v}_k.$$

- If  $|\lambda_k|$  is small then the component of  $\mathbf{u} - \mathbf{u}^{(j)}$  in the direction of  $\mathbf{v}_k$  vanishes quickly.
- After only a few iterations, the error is dominated by those components in direction  $\mathbf{v}_k$ , where  $|\lambda_k| \approx 1$ .
- The eigenvectors of  $\mathbf{R}(\omega)$  with  $n = 50$ . From left to right:  $\mathbf{v}_1$ ,  $\mathbf{v}_{25}$ , and  $\mathbf{v}_{50}$ . Each graph shows the points  $(ih, (\mathbf{v}_k)_i)$  for  $1 \leq i \leq n$  linearly connected.



Low-frequency ( $k \leq n/2$ ), high-frequency ( $k > n/2$ )

- Since  $|\lambda_1| = |\lambda_n| \approx 1$ , the error in direction  $\mathbf{v}_1$  and direction  $\mathbf{v}_n$  is large, which means no matter how many steps in Jacobi's method we compute, the error will always contain both low-frequency and high-frequency eigenvectors.
- To avoid this, let us have another look at the relaxation of Jacobi's method. Choosing  $\omega = 1/2$  yields the eigenvalues

$$\lambda_k(1/2) = (1 + \cos(k\pi h))/2, \quad 1 \leq k \leq n.$$

For large  $k$  this means that  $\lambda_k(1/2)$  is very close to zero, while for small  $k$  we have  $\lambda_k(1/2)$  very close to 1.

- Consider the error after  $j$  iterations

$$\mathbf{u} - \mathbf{u}^{(j)} = [\mathbf{R}(1/2)]^j (\mathbf{u} - \mathbf{u}^{(0)}) = \sum_{k=1}^n \alpha_k [\lambda_k(1/2)]^j \mathbf{v}_k.$$

The low-frequency eigenvectors dominate and the influence of the high-frequency eigenvectors tends to zero.



- The error, in a certain way, is “smoothed” during the process.
- A “smoother” error can be represented using a smaller  $n$  and this gives the idea of the two-grid method, as follows.
  - Compute  $j$  steps of the relaxation, resulting in an error

$$\boldsymbol{\varepsilon}^{(j)} = \mathbf{u} - \mathbf{u}^{(j)} = \mathbf{R}(\omega)^j \boldsymbol{\varepsilon}^{(0)}$$

which is much “smoother” than  $\boldsymbol{\varepsilon}^{(0)}$ .

- We have  $\mathbf{u} = \mathbf{u}^{(j)} + \boldsymbol{\varepsilon}^{(j)}$  and  $\boldsymbol{\varepsilon}^{(j)}$  satisfies

$$\mathbf{A}\boldsymbol{\varepsilon}^{(j)} = \mathbf{A}(\mathbf{u} - \mathbf{u}^{(j)}) = \mathbf{f} - \mathbf{A}\mathbf{u}^{(j)} =: \mathbf{r}^{(j)}.$$

Hence, if we can solve  $\mathbf{A}\boldsymbol{\varepsilon}^{(j)} = \mathbf{r}^{(j)}$  then the overall solution is given by  $\mathbf{u} = \mathbf{u}^{(j)} + \boldsymbol{\varepsilon}^{(j)}$ .

- Since we expect the error  $\boldsymbol{\varepsilon}^{(j)}$  to be “smooth”, we will solve the equation  $\mathbf{A}\boldsymbol{\varepsilon}^{(j)} = \mathbf{r}^{(j)}$  somehow on a coarser grid to save computational time and transfer the solution back to the finer grid.

### 3. Two-grid, V-cycle, and Multigrid

- Assume that we are given two grids: a fine grid  $X_h$  with  $n_h$  points and a coarse grid  $X_H$  with  $n_H < n_h$  points. Associated with these grids are discrete solution spaces  $V_h = \mathbb{R}^{n_h}$  and  $V_H = \mathbb{R}^{n_H}$ .
- We need a *prolongation* operator  $\mathbf{I}_H^h : V_H \mapsto V_h$  which maps from coarse to fine, and a *restriction* operator  $\mathbf{I}_h^H : V_h \mapsto V_H$  which maps from fine to coarse.
- For simplicity, suppose the coarse grid is given by

$$X_H = \{jH : 0 \leq j \leq n_H - 1\}$$

with  $n_H = 2^m + 1$ ,  $m \in \mathbb{N}$ , and  $H = 1/(n_H - 1)$ . Then the natural fine grid  $X_h$  would consist of  $X_H$  and all points in the middle between two points from  $X_H$ , i.e.,

$$X_h = \{jh : 0 \leq j \leq n_h - 1\}$$

with  $h = H/2$  and  $n_h = 2^{m+1} + 1$ .

- In this case we could define the prolongation and restriction operators as follows. The prolongation  $\mathbf{v}^h = \mathbf{I}_H^h \mathbf{v}^H$  is defined by linear interpolation on the “in-between” points:

$$\begin{aligned} v_{2j}^h &:= v_j^H, & 0 \leq j \leq n_H - 1; \\ v_{2j+1}^h &:= \frac{v_j^H + v_{j+1}^H}{2}, & 0 \leq j \leq n_H - 2. \end{aligned}$$

In matrix form we have

$$\mathbf{I}_H^h = \frac{1}{2} \begin{bmatrix} 2 & & & & & \\ 1 & 1 & & & & \\ & 2 & & & & \\ & 1 & 1 & & & \\ & & \vdots & \vdots & \vdots & \\ & & & 1 & 1 & \\ & & & & 2 & \end{bmatrix}.$$

- For the restriction,  $\mathbf{v}^H = \mathbf{I}_h^H \mathbf{v}^h$  we could use the natural inclusion, i.e., we could simply define  $v_j^H := v_{2j}^h$ ,  $0 \leq j \leq n_H - 1$ .

We could, however, also use a so-called full weighting, which is given by

$$v_j^H = \frac{1}{4}(v_{2j-1}^h + 2v_{2j}^h + v_{2j+1}^h), \quad 0 \leq j \leq n_H - 1,$$

where we have implicitly set  $v_{-1}^h = v_{n_h}^h = 0$ . In matrix form (for the full weighting case) we have

$$\mathbf{I}_h^H = \frac{1}{4} \begin{bmatrix} 2 & 1 & & & & \\ & 1 & 2 & 1 & & \\ & & & 1 & \cdots & \\ & & & & \cdots & \\ & & & & \cdots & 1 \\ & & & & & 1 & 2 \end{bmatrix}.$$

For this case, we have  $\mathbf{I}_H^h = 2(\mathbf{I}_h^H)^\top$ .

- Our goal is to solve the finite difference system  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$  on the fine level, using the possibility of solving a system  $\mathbf{A}_H \boldsymbol{\varepsilon}^H = \mathbf{r}^H$  on a coarse level.

Note that the matrix  $\mathbf{A}_h$  and  $\mathbf{A}_H$  only refer to interior nodes.

Hence, we delete the first and last columns and rows in the matrix representation of  $\mathbf{I}_H^h$  and  $\mathbf{I}_h^H$ . We still use  $\mathbf{I}_H^h$  and  $\mathbf{I}_h^H$  to denote the resulting matrices, i.e.,

$$\mathbf{I}_H^h = \frac{1}{2} \begin{bmatrix} 1 & & & & \\ 2 & & & & \\ 1 & 1 & & & \\ & \vdots & \vdots & \vdots & \\ & & & & 1 \end{bmatrix}, \quad \mathbf{I}_h^H = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 & \cdots & \\ & & 1 & \cdots & \\ & & & \cdots & \\ & & & & \cdots & 1 \end{bmatrix}.$$

We can prove that (Exercise)

$$\mathbf{A}_H = \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h.$$

- We will use a stationary iterative method as a “smoother”. Recall that such an iterative method for solving  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$  is given by

$$\mathbf{u}_{(j+1)}^h = \mathbf{S}_h(\mathbf{u}_{(j)}^h) := \mathbf{R}_h \mathbf{u}_{(j)}^h + \mathbf{c}^h,$$

where the solution  $\mathbf{u}^h$  of the linear system is a fixed point of  $\mathbf{S}_h(\cdot)$ , i.e., it satisfies

$$\mathbf{u}^h = \mathbf{S}_h(\mathbf{u}^h) := \mathbf{R}_h \mathbf{u}^h + \mathbf{c}^h.$$

Sometimes we call it a consistent iterative method.

- If we apply  $\ell \in \mathbb{N}$  iterations of such a smoother with initial data  $\mathbf{u}_{(0)}^h$ , it is easy to see that the result has the form

$$\mathbf{S}_h^\ell(\mathbf{u}_{(0)}^h) = \mathbf{R}_h^\ell \mathbf{u}_{(0)}^h + \sum_{j=0}^{\ell-1} \mathbf{R}_h^j \mathbf{c}^h := \mathbf{R}_h^\ell \mathbf{u}_{(0)}^h + \mathbf{s}_{(\ell)}^h.$$

- Note that we can use any consistent method as the smoother  $\mathbf{S}_h(\cdot)$ .

---

**Algorithm:** Two-grid for  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$ ,  $\text{TG}(\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2)$

---

**Input:**  $\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2 \in \mathbb{N}$ .

**Output:**  $\mathbf{u}^h$ , an approximation to  $\mathbf{A}_h^{-1} \mathbf{f}^h$ .

1. Presmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(0)}^h)$
  2. Get residual :  $\mathbf{r}^h := \mathbf{f}^h - \mathbf{A}_h \mathbf{u}^h$
  3. Coarsen :  $\mathbf{r}^H := \mathbf{I}_h^H \mathbf{r}^h$
  4. Solve :  $\boldsymbol{\varepsilon}^H := \mathbf{A}_H^{-1} \mathbf{r}^H$
  5. Prolong :  $\boldsymbol{\varepsilon}^h := \mathbf{I}_H^h \boldsymbol{\varepsilon}^H$
  6. Correct :  $\mathbf{u}^h := \mathbf{u}^h + \boldsymbol{\varepsilon}^h$
  7. Postsmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$
-

- The two-grid method can be seen as only one update of a new stationary iterative method.

## Theorem 1

*Assume that  $\mathbf{A}_H$  is invertible. Assume that  $\mathbf{u}_{(j)}^h$  is the input vector and  $\mathbf{u}_{(j+1)}^h$  is the resulting output vector of the two-grid method. Then, we have*

$$\mathbf{u}_{(j+1)}^h = \mathbf{T}_h \mathbf{u}_{(j)}^h + \mathbf{d}^h \quad \text{and} \quad \mathbf{u}^h = \mathbf{T}_h \mathbf{u}^h + \mathbf{d}^h,$$

*where the iteration matrix  $\mathbf{T}_h$  is given by*

$$\mathbf{T}_h = \mathbf{R}_h^{\ell_2} \mathbf{T}_{h,H} \mathbf{R}_h^{\ell_1} \quad \text{with} \quad \mathbf{T}_{h,H} = \mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h.$$

*Moreover, we have the error representation*

$$\mathbf{u}_{(j+1)}^h - \mathbf{u}^h = \mathbf{T}_h (\mathbf{u}_{(j)}^h - \mathbf{u}^h) = \mathbf{T}_h^{j+1} (\mathbf{u}_{(0)}^h - \mathbf{u}^h),$$

*showing that the method converges if the spectral radius  $\rho(\mathbf{T}_h) < 1$ .*



*Proof.* We go through the two-grid method step by step. With the first step  $\mathbf{u}_{(j)}^h$  is mapped to  $\mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h)$ , which is the input to the second step. After the second and third steps we have

$$\mathbf{r}^H = \mathbf{I}_h^H(\mathbf{f}^h - \mathbf{A}_h \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h)),$$

which is the input for the fourth step, so that the results after the fourth and fifth steps become

$$\boldsymbol{\varepsilon}^h = \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H(\mathbf{f}^h - \mathbf{A}_h \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h)).$$

Applying steps 6 and 7 to this finally results in the new iteration

$$\begin{aligned} \mathbf{u}_{(j+1)}^h &= \mathbf{S}_h^{\ell_2} \left( \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h) + \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H(\mathbf{f}^h - \mathbf{A}_h \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h)) \right) \\ &= \mathbf{S}_h^{\ell_2} \left( (\mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h) \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h) + \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h \right) \\ &= \mathbf{S}_h^{\ell_2} \left( \mathbf{T}_{h,H} \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(j)}^h) + \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h \right). \end{aligned}$$

Note that the operator  $\mathbf{S}_h(\cdot)$  is only affine and not linear. Define

$$\tilde{\mathbf{T}}_h(\cdot) := \mathbf{S}_h^{\ell_2}(\mathbf{T}_{h,H} \mathbf{S}_h^{\ell_1}(\cdot)), \quad \tilde{\mathbf{d}}^h := \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h.$$

By straightforward calculations, we have

$$\mathbf{u}_{(j+1)}^h = \tilde{\mathbf{T}}_h(\mathbf{u}_{(j)}^h) + \tilde{\mathbf{d}}^h,$$

and

$$\begin{aligned} \tilde{\mathbf{T}}_h(\mathbf{u}) &= \mathbf{S}_h^{\ell_2}(\mathbf{T}_{h,H} \mathbf{S}_h^{\ell_1}(\mathbf{u})) \\ &= \mathbf{S}_h^{\ell_2} \left( \mathbf{T}_{h,H} \left( \mathbf{R}_h^{\ell_1} \mathbf{u} + \sum_{j=0}^{\ell_1-1} \mathbf{R}_h^j \mathbf{c}^h \right) \right) \\ &= \mathbf{R}_h^{\ell_2} \mathbf{T}_{h,H} \mathbf{R}_h^{\ell_1} \mathbf{u} + \mathbf{R}_h^{\ell_2} \mathbf{T}_{h,H} \sum_{j=0}^{\ell_1-1} \mathbf{R}_h^j \mathbf{c}^h + \sum_{j=0}^{\ell_2-1} \mathbf{R}_h^j \mathbf{c}^h \\ &=: \mathbf{T}_h \mathbf{u} + \hat{\mathbf{d}}^h, \end{aligned}$$

which shows

$$\mathbf{u}_{(j+1)}^h = \mathbf{T}_h \mathbf{u}_{(j)}^h + \mathbf{d}^h \quad \text{with} \quad \mathbf{d}^h = \widehat{\mathbf{d}}^h + \widetilde{\mathbf{d}}^h.$$

Hence, the iteration matrix is indeed given by

$$\mathbf{T}_h = \mathbf{R}_h^{\ell_2} \mathbf{T}_{h,H} \mathbf{R}_h^{\ell_1}.$$

As  $\mathbf{S}_h(\cdot)$  is consistent, i.e.,  $\mathbf{S}_h(\mathbf{u}^h) = \mathbf{u}^h$ , we have

$$\begin{aligned} \mathbf{T}_h \mathbf{u}^h + \mathbf{d}^h &= \mathbf{T}_h \mathbf{u}^h + \widehat{\mathbf{d}}^h + \widetilde{\mathbf{d}}^h = \widetilde{\mathbf{T}}_h(\mathbf{u}^h) + \widetilde{\mathbf{d}}^h \\ &= \mathbf{S}_h^{\ell_2}(\mathbf{T}_{h,H} \mathbf{S}_h^{\ell_1}(\mathbf{u}^h)) + \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h \\ &= \mathbf{S}_h^{\ell_2}((\mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h) \mathbf{u}^h) + \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h \\ &= \mathbf{S}_h^{\ell_2}(\mathbf{u}^h) - \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h + \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{f}^h \\ &= \mathbf{u}^h. \end{aligned}$$

This shows

$$\mathbf{u}_{(j+1)}^h - \mathbf{u}^h = \mathbf{T}_h(\mathbf{u}_{(j)}^h - \mathbf{u}^h) = \mathbf{T}_h^{j+1}(\mathbf{u}_{(0)}^h - \mathbf{u}^h). \quad \square$$

## Proposition 2

Assume the following four conditions hold.

- (1) The matrix  $\mathbf{A}_h$  is symmetric and positive definite.
- (2) The prolongation and restriction operators are connected by

$$\mathbf{I}_H^h = \gamma(\mathbf{I}_h^H)^\top$$

with  $\gamma > 0$ .

- (3) The prolongation operator  $\mathbf{I}_H^h$  is injective.
- (4) The coarse grid matrix is given by  $\mathbf{A}_H := \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h$ .

Then we have:

- (i) The coarse-grid correction operator  $\mathbf{T}_{h,H}$  is an orthogonal projector with respect to the inner product  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ .
- (ii) The range of  $\mathbf{T}_{h,H}$  is  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ -orthogonal to the range of  $\mathbf{I}_H^h$ .

*Proof.* We start by showing that  $\mathbf{A}_H$  is symmetric and positive definite. It is symmetric since

$$\mathbf{A}_H^\top = (\mathbf{I}_H^h)^\top \mathbf{A}_h^\top (\mathbf{I}_H^H)^\top = \gamma \mathbf{I}_h^H \mathbf{A}_h \frac{1}{\gamma} \mathbf{I}_H^h = \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h = \mathbf{A}_H.$$

It is positive definite since we have for any  $\mathbf{x} \neq \mathbf{0}$  that  $\mathbf{I}_H^h \mathbf{x} \neq \mathbf{0}$  because of the injectivity of  $\mathbf{I}_H^h$  and hence

$$\mathbf{x}^\top \mathbf{A}_H \mathbf{x} = \mathbf{x}^\top \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h \mathbf{x} = \frac{1}{\gamma} (\mathbf{I}_H^h \mathbf{x})^\top \mathbf{A}_h (\mathbf{I}_H^h \mathbf{x}) > 0.$$

This means in particular that the coarse grid correction operator

$$\mathbf{T}_{h,H} = \mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h$$

is well-defined. Next, let

$$\mathbf{Q}_{h,H} := \mathbf{I} - \mathbf{T}_{h,H} = \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h.$$

Actually, the mapping  $\mathbf{Q}_{h,H}$  is a projection since we have

$$\begin{aligned}\mathbf{Q}_{h,H}^2 &= (\mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h)(\mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h) \\ &= \mathbf{I}_H^h \mathbf{A}_H^{-1} (\mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h) \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h \\ &= \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h \\ &= \mathbf{Q}_{h,H}.\end{aligned}$$

It is also self-adjoint and hence an orthogonal projector:

$$\begin{aligned}\langle \mathbf{Q}_{h,H} \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}_h} &= \mathbf{x}^\top \mathbf{Q}_{h,H}^\top \mathbf{A}_h \mathbf{y} = \mathbf{x}^\top \mathbf{A}_h^\top (\mathbf{I}_h^H)^\top (\mathbf{A}_H^{-1})^\top (\mathbf{I}_H^h)^\top \mathbf{A}_h \mathbf{y} \\ &= \mathbf{x}^\top \mathbf{A}_h \left( \frac{1}{\gamma} \mathbf{I}_H^h \right) \mathbf{A}_H^{-1} \gamma \mathbf{I}_h^H \mathbf{A}_h \mathbf{y} \\ &= \mathbf{x}^\top \mathbf{A}_h \mathbf{Q}_{h,H} \mathbf{y} \\ &= \langle \mathbf{x}, \mathbf{Q}_{h,H} \mathbf{y} \rangle_{\mathbf{A}_h}.\end{aligned}$$

Then  $\mathbf{T}_{h,H}$  is also an orthogonal projector with respect to  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ .

It remains to show that  $\text{range}(\mathbf{T}_{h,H})$  is  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ -orthogonal to  $\text{range}(\mathbf{I}_H^h)$ . This is true because we have

$$\begin{aligned}
 \langle \mathbf{T}_{h,H} \mathbf{x}, \mathbf{I}_H^h \mathbf{y} \rangle_{\mathbf{A}_h} &= \mathbf{x}^\top (\mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_H^H \mathbf{A}_h)^\top \mathbf{A}_h \mathbf{I}_H^h \mathbf{y} \\
 &= \mathbf{x}^\top (\mathbf{I} - \mathbf{A}_h \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_H^H) \mathbf{A}_h \mathbf{I}_H^h \mathbf{y} \\
 &= \mathbf{x}^\top (\mathbf{A}_h \mathbf{I}_H^h - \mathbf{A}_h \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_H^H \mathbf{A}_h \mathbf{I}_H^h) \mathbf{y} \\
 &= \mathbf{x}^\top (\mathbf{A}_h \mathbf{I}_H^h - \mathbf{A}_h \mathbf{I}_H^h) \mathbf{y} \\
 &= 0
 \end{aligned}$$

for all  $\mathbf{x}$  and  $\mathbf{y}$ . □

- Let  $\mathbf{D}_h$  be the diagonal part of  $\mathbf{A}_h$ . We say that  $\mathbf{S}_h(\cdot)$  has the *smoothing property* if there is a constant  $\alpha > 0$  such that

$$\|\mathbf{R}_h \mathbf{v}^h\|_{\mathbf{A}_h}^2 \leq \|\mathbf{v}^h\|_{\mathbf{A}_h}^2 - \alpha \|\mathbf{A}_h \mathbf{v}^h\|_{\mathbf{D}_h^{-1}}^2, \quad \forall \mathbf{v}^h.$$

We say that the prolongation operator  $\mathbf{I}_H^h$  has the *approximation property* if there is a constant  $\beta > 0$  such that

$$\min_{\mathbf{v}^H} \|\mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H\|_{\mathbf{D}_h} \leq \beta \|\mathbf{A}_h \mathbf{v}^h\|_{\mathbf{D}_h^{-1}}, \quad \forall \mathbf{v}^h.$$

### Theorem 3

*Let the conditions of Proposition 2 be satisfied. Assume:*

- (1) the smoothing process  $\mathbf{S}_h(\cdot)$  has the smoothing property,*
- (2) the prolongation operator  $\mathbf{I}_H^h$  has the approximation property.*

*Then, we have*

$$\alpha \leq \beta$$

*and for the iteration matrix  $\mathbf{T}_h$  of the two-grid method,*

$$\|\mathbf{T}_h\|_{\mathbf{A}_h} < \sqrt{1 - \alpha/\beta} \text{ if } \alpha < \beta \text{ and } \|\mathbf{T}_h\|_{\mathbf{A}_h} = 0 \text{ if } \alpha = \beta.$$

*Hence, as an iterative method, the two-grid method converges.*

*Proof.* Since  $\text{range}(\mathbf{T}_{h,H})$  is  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ -orthogonal to  $\text{range}(\mathbf{I}_H^h)$ , we have

$$\langle \mathbf{v}^h, \mathbf{I}_H^h \mathbf{v}^H \rangle_{\mathbf{A}_h} = 0, \quad \forall \mathbf{v}^h \in \text{range}(\mathbf{T}_{h,H}), \quad \forall \mathbf{v}^H.$$



From this, we can conclude for all  $\mathbf{v}^h \in \text{range}(\mathbf{T}_{h,H})$  and  $\mathbf{v}^H$  that

$$\begin{aligned}\|\mathbf{v}^h\|_{\mathbf{A}_h}^2 &= \langle \mathbf{v}^h, \mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H \rangle_{\mathbf{A}_h} = \langle \mathbf{A}_h \mathbf{v}^h, \mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H \rangle \\ &= \langle \mathbf{D}_h^{-1/2} \mathbf{A}_h \mathbf{v}^h, \mathbf{D}_h^{1/2} (\mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H) \rangle \\ &\leq \|\mathbf{D}_h^{-1/2} \mathbf{A}_h \mathbf{v}^h\|_2 \|\mathbf{D}_h^{1/2} (\mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H)\|_2 \\ &= \|\mathbf{A}_h \mathbf{v}^h\|_{\mathbf{D}_h^{-1}} \|\mathbf{v}^h - \mathbf{I}_H^h \mathbf{v}^H\|_{\mathbf{D}_h}.\end{aligned}$$

Going over to the infimum over all  $\mathbf{v}^H$  and using the approximation property leads to

$$\|\mathbf{v}^h\|_{\mathbf{A}_h} \leq \sqrt{\beta} \|\mathbf{A}_h \mathbf{v}^h\|_{\mathbf{D}_h^{-1}}, \quad \forall \mathbf{v}^h \in \text{range}(\mathbf{T}_{h,H}).$$

This is equivalent to

$$\|\mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h} \leq \sqrt{\beta} \|\mathbf{A}_h \mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{D}_h^{-1}}, \quad \forall \mathbf{v}^h.$$

Using this and the smoothing property then leads to

$$\begin{aligned}
 0 \leq \|\mathbf{R}_h \mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h}^2 &\leq \|\mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h}^2 - \alpha \|\mathbf{A}_h \mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{D}_h^{-1}}^2 \\
 &\leq \|\mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h}^2 - \frac{\alpha}{\beta} \|\mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h}^2 \\
 &= (1 - \alpha/\beta) \|\mathbf{T}_{h,H} \mathbf{v}^h\|_{\mathbf{A}_h}^2 \\
 &\leq (1 - \alpha/\beta) \|\mathbf{v}^h\|_{\mathbf{A}_h}^2,
 \end{aligned}$$

where the last inequality follows from the fact that  $\mathbf{T}_{h,H}$  is an orthogonal projector with respect to  $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$ . This means, first of all,  $\alpha \leq \beta$  and secondly

$$\|\mathbf{R}_h \mathbf{T}_{h,H}\|_{\mathbf{A}_h} \leq \sqrt{1 - \alpha/\beta}.$$

By the smoothing property, we have  $\|\mathbf{R}_h\|_{\mathbf{A}_h} < 1$ . Then we finally derive  $\|\mathbf{T}_h\|_{\mathbf{A}_h} = 0$  if  $\alpha = \beta$ , and if  $\alpha < \beta$ ,

$$\|\mathbf{T}_h\|_{\mathbf{A}_h} = \|\mathbf{R}_h^{\ell_2} \mathbf{T}_{h,H} \mathbf{R}_h^{\ell_1}\|_{\mathbf{A}_h} < \|\mathbf{R}_h \mathbf{T}_{h,H}\|_{\mathbf{A}_h} \leq \sqrt{1 - \alpha/\beta}. \quad \square$$

---

**Algorithm:** V-cycle for  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$ ,  $\text{V-cycle}(\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2, h_0)$

---

**Input:**  $\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2 \in \mathbb{N}, h_0$ .

**Output:**  $\mathbf{u}^h$ , an approximation to  $\mathbf{A}_h^{-1} \mathbf{f}^h$ .

1. Presmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(0)}^h)$
2. Get residual :  $\mathbf{r}^h := \mathbf{f}^h - \mathbf{A}_h \mathbf{u}^h$
3. Coarsen :  $H := 2h, \quad \mathbf{r}^H := \mathbf{I}_h^H \mathbf{r}^h$
4. if  $H = h_0$   
    Solve  $\mathbf{A}_H \boldsymbol{\epsilon}^H = \mathbf{r}^H$   
  else  
     $\boldsymbol{\epsilon}^H := \text{V-cycle}(\mathbf{A}_H, \mathbf{r}^H, \mathbf{0}, \ell_1, \ell_2, h_0)$   
  end
5. Prolong :  $\boldsymbol{\epsilon}^h := \mathbf{I}_H^h \boldsymbol{\epsilon}^H$
6. Correct :  $\mathbf{u}^h := \mathbf{u}^h + \boldsymbol{\epsilon}^h$
7. Postsmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$

---

**Algorithm:** Multigrid for  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$ ,  $\text{MG}(\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2, \ell, h_0)$

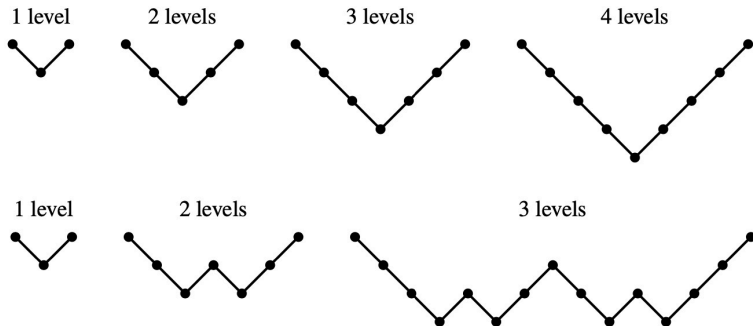
---

**Input:**  $\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}_{(0)}^h, \ell_1, \ell_2, \ell \in \mathbb{N}, h_0$ .

**Output:**  $\mathbf{u}^h$ , an approximation to  $\mathbf{A}_h^{-1} \mathbf{f}^h$ .

1. Presmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}_{(0)}^h)$
  2. Get residual :  $\mathbf{r}^h := \mathbf{f}^h - \mathbf{A}_h \mathbf{u}^h$
  3. Coarsen :  $H := 2h, \quad \mathbf{r}^H := \mathbf{I}_h^H \mathbf{r}^h$
  4. if  $H = h_0$   
    Solve  $\mathbf{A}_H \boldsymbol{\epsilon}^H = \mathbf{r}^H$   
    else  
         $\boldsymbol{\epsilon}^H := \mathbf{0}$   
        for  $j = 1 : \ell$   
             $\boldsymbol{\epsilon}^H := \text{MG}(\mathbf{A}_H, \mathbf{r}^H, \boldsymbol{\epsilon}^H, \ell_1, \ell_2, \ell, h_0)$   
        end  
    end  
5. Prolong :  $\boldsymbol{\epsilon}^h := \mathbf{I}_H^h \boldsymbol{\epsilon}^H$   
6. Correct :  $\mathbf{u}^h := \mathbf{u}^h + \boldsymbol{\epsilon}^h$   
7. Postsmooth :  $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$
-

- It is helpful to visualize the recursion, depending on the choice of  $\ell$  and how many levels there are, meaning how many grids we use. Assume that  $k$  levels are used, i.e.,  $k$  satisfies  $h_0 = 2^k h$ .



The recursion of multigrid with  $\ell = 1$  (top, V-cycle) and  $\ell = 2$  (bottom, W-cycle).

- To determine the iteration matrix  $\mathbf{M}_h$  of the multigrid method, we start with the iteration matrix of the two-grid method

$$\mathbf{T}_h = \mathbf{R}_h^{\ell_2} (\mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h) \mathbf{R}_h^{\ell_1}$$

and recall that the term  $\mathbf{A}_H^{-1}$  came from step 4 in two-grid cycle and hence has now to be replaced by  $\ell$  steps of the multigrid method on grid  $X_H$ :

$$\begin{aligned}\boldsymbol{\varepsilon}_{(0)}^H &:= \mathbf{0}, \\ \boldsymbol{\varepsilon}_{(j)}^H &:= \mathbf{M}_H \boldsymbol{\varepsilon}_{(j-1)}^H + \mathbf{d}^H, \quad 1 \leq j \leq \ell.\end{aligned}$$

As this is a consistent method for solving  $\mathbf{A}_H \boldsymbol{\varepsilon}^H = \mathbf{r}^H$ , we have

$$\boldsymbol{\varepsilon}_{(\ell)}^H - \boldsymbol{\varepsilon}^H = \mathbf{M}_H (\boldsymbol{\varepsilon}_{(\ell-1)}^H - \boldsymbol{\varepsilon}^H) = \mathbf{M}_H^\ell (\boldsymbol{\varepsilon}_{(0)}^H - \boldsymbol{\varepsilon}^H) = -\mathbf{M}_H^\ell \boldsymbol{\varepsilon}^H.$$

Then, we have

$$\boldsymbol{\varepsilon}_{(\ell)}^H = (\mathbf{I} - \mathbf{M}_H^\ell) \boldsymbol{\varepsilon}^H = (\mathbf{I} - \mathbf{M}_H^\ell) \mathbf{A}_H^{-1} \mathbf{r}^H.$$

Replacing  $\mathbf{A}_H^{-1}$  by  $(\mathbf{I} - \mathbf{M}_H^\ell)\mathbf{A}_H^{-1}$  yields

$$\begin{aligned}\mathbf{M}_h &= \mathbf{R}_h^{\ell_2} [\mathbf{I} - \mathbf{I}_H^h (\mathbf{I} - \mathbf{M}_H^\ell) \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h] \mathbf{R}_h^{\ell_1} \\ &= \mathbf{R}_h^{\ell_2} [\mathbf{I} - \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h + \mathbf{I}_H^h \mathbf{M}_H^\ell \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h] \mathbf{R}_h^{\ell_1} \\ &= \mathbf{T}_h + \mathbf{R}_h^{\ell_2} \mathbf{I}_H^h \mathbf{M}_H^\ell \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h \mathbf{R}_h^{\ell_1},\end{aligned}$$

which can be seen as a perturbation of  $\mathbf{T}_h$ .

- *Algebraic multigrid* (AMG):  $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$  with  $\mathbf{A}_h \in \mathbb{R}^{n_h \times n_h}$ 
  - define the coarse subset  $\mathbb{R}^{n_H}$  from the fine set  $\mathbb{R}^{n_h}$ ,
  - define the coarsening operator  $\mathbf{I}_h^H$  from  $\mathbb{R}^{n_h}$  to  $\mathbb{R}^{n_H}$
  - use the abstract definitions

$$\mathbf{I}_H^h = (\mathbf{I}_h^H)^\top, \quad \mathbf{A}_H = \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h, \quad \mathbf{f}^H = \mathbf{I}_h^H \mathbf{f}^h$$

to complete the set-up.

## 4. Further reading

- Holger Wendland

Numerical Linear Algebra An Introduction

Cambridge University Press, 2018

- William L. Briggs, Van E. Henson, and Steve F. McCormick

A Multigrid Tutorial

Second Edition, SIAM, 2000

- James W. Demmel

Applied Numerical Linear Algebra (Section 6.9, V-cycle, full MG)

SIAM, 1997

- Pieter Wesseling

An Introduction to Multigrid Methods

John Wiley & Sons, 1992