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

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

- Let $u_i^h := u(x_i)$ and $f_i^h := f(x_i)$, where

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

- The finite difference method is: find $\mathbf{u}^h$ such that

$$u_0^h = u_{n+1}^h = 0$$

and

$$-\left(\frac{u_{i+1}^h - 2u_i^h + u_{i-1}^h}{h^2}\right) = f_i^h, \quad 1 \le i \le n.$$

- The FD system

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

2. Classical stationary iterative methods (Lecture 6)

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

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

  - Jacobi's method
  - Gauss–Seidel method
  - Successive overrelaxation: SOR(ω)
  - Symmetric successive overrelaxation: SSOR(ω)

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

- The iteration matrix

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

$$= \begin{bmatrix} 0 & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{2} & 0 & \frac{1}{2} \\ & & & \frac{1}{2} & 0 \end{bmatrix}.$$

- The relaxation of Jacobi's method

$$\mathbf{R}(\omega) = (1-\omega)\mathbf{I} + \omega\mathbf{R} = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}.$$

- 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

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

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

- The eigenvalues of the relaxation of the Jacobi method are given by

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

Note that relaxation does not lead to an improved convergence, since, in this case, the optimal relaxation parameter is $\omega_\star = 1$. (why?)

2.2. What makes the convergence 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

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

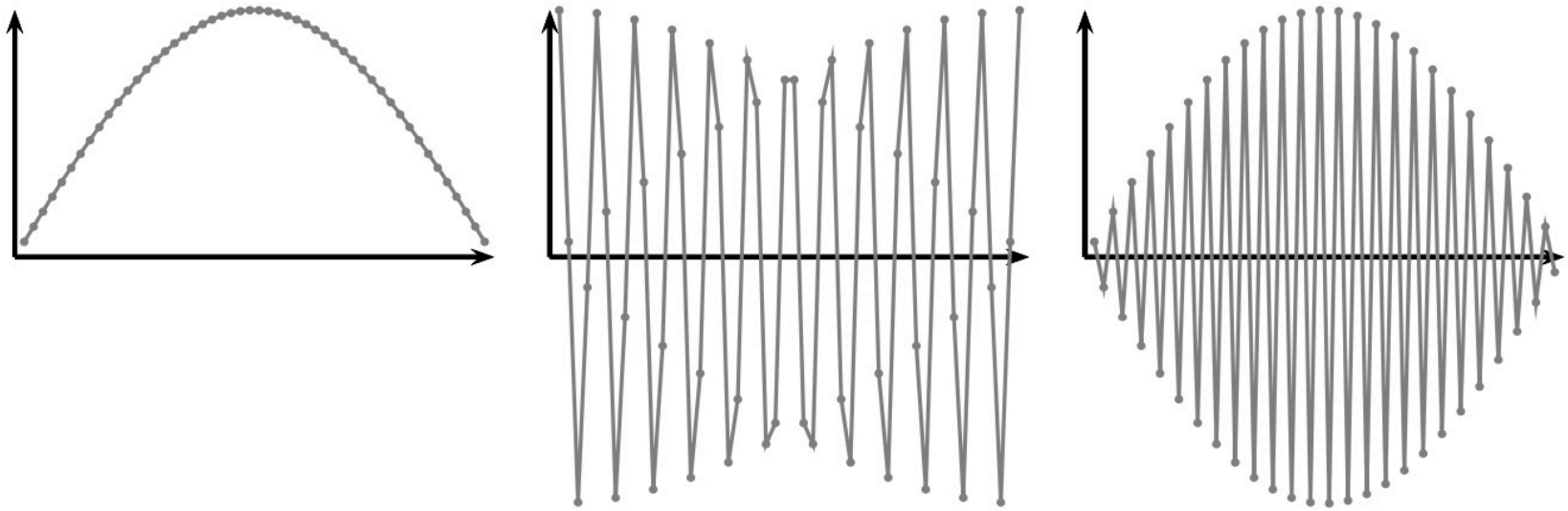
We expand $\mathbf{u} - \mathbf{u}^{(0)}$ in the basis consisting of the eigenvectors:

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

This gives

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

- 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 ($k \leq n/2$) and high-frequency ($k > n/2$)
- 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 $|\lambda_k| \approx 1$.

- In particular, the error in direction $\mathbf{v}_1$ and direction $\mathbf{v}_n$ is large, which means no matter how many steps in the Jacobi 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 the Jacobi 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

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

  which is much "smoother" than $\boldsymbol{\varepsilon}^{(0)}$.
  - We have $\mathbf{u} = \mathbf{u}^{(j)} + \boldsymbol{\varepsilon}^{(j)}$ and $\boldsymbol{\varepsilon}^{(j)}$ satisfies

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

  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 we need a *restriction* operator $\mathbf{I}_h^H : V_h \mapsto V_H$ which maps from fine to coarse.

- Suppose the coarse grid is given by

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

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.,

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

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:

$$v_{2j}^h := v_j^H, \quad v_{2j+1}^h := \frac{v_j^H + v_{j+1}^H}{2}, \quad 0 \le j < n_H.$$

  In matrix form we have

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

- 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 $v_j^H := v_{2j}^h$, $0 \le j < n_H$.

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

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

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

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

- Our goal is to solve a system $\mathbf{A}_h \mathbf{u}^h = \mathbf{f}^h$ on the fine level, using the possibility of solving a system $\mathbf{A}_H \mathbf{u}^H = \mathbf{f}^H$ on a coarse level.
- Hence, we are usually given the equation on the fine level but need $\mathbf{A}_H$ and $\mathbf{f}^H$ on the coarse level.
- These are often given quite naturally, simply by discretizing the original problem on the coarse grid X_H. However, we can also use the prolongation and restriction operators to define

$$\mathbf{A}_H := \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h, \qquad \mathbf{f}^H := \mathbf{I}_h^H \mathbf{f}^h.$$

- We indeed find that a discretization on the coarse grid leads to the same matrix $\mathbf{A}_H$ as above. (Exercise)

  To see this, we have to take into account that the matrix $\mathbf{A}_h$ only refers to interior nodes. Hence, we either have to delete the first and last column and row in the matrix representation of $\mathbf{I}_H^h$ and $\mathbf{I}_h^H$, or we only look at interior points.

- As mentioned above, we will use an iterative method as a "smoother". Recall that such a consistent iterative method for solving $\mathbf{A}_h\mathbf{u}^h = \mathbf{f}^h$ is given by

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

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

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

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

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

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

Algorithm: Two-grid cycle for $\mathbf{A}_h\mathbf{u}^h = \mathbf{f}^h$

Input: $\mathbf{A}_h \in \mathbb{R}^{n_h \times n_h}$, $\mathbf{f}^h \in \mathbb{R}^{n_h}$, $\mathbf{u}^h_{(0)} \in \mathbb{R}^{n_h}$, $\ell_1, \ell_2 \in \mathbb{N}$.

Output: Approximation to $\mathbf{A}_h^{-1}\mathbf{f}^h$.

1. `Presmooth` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}^h_{(0)})$
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` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$

- Obviously, the two-grid method can be and usually is seen as only one cycle of a new iterative method.

  To analyze it, we need to express this method in the typical form of an iterative method and then need to analyze the iteration matrix.

  To this end, let us assume that $\mathbf{u}^h_{(j)}$ is the input vector and $\mathbf{u}^h_{(j+1)}$ is the resulting output vector. Then, we have to find the iteration matrix $\mathbf{T}_h \in \mathbb{R}^{n_h \times n_h}$ and a vector $\mathbf{d}^h$ such that

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

  and we have to convince ourselves that the method is consistent, i.e.,

$$\mathbf{u}^h = \mathbf{T}_h \mathbf{u}^h + \mathbf{d}^h.$$

- Going through the two-grid method step by step leads to the following.

- 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

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

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

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

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

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

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

Lemma 1

Let the consistent smoother $\mathbf{S}_h$ *be given as*

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

Assume that $\mathbf{A}_H$ *is invertible. Then, the iteration matrix* $\mathbf{T}_h$ *of the two-grid method is given by*

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

with the coarse-grid correction operator

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

Moreover, if $\mathbf{u}^h$ *solves* $\mathbf{A}_h\mathbf{u}^h = \mathbf{f}^h$ *then 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 $\rho(\mathbf{T}_h) < 1$.

Proposition 2

Assume the following four conditions hold.

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

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

  with $\gamma > 0$.
- *The prolongation operator* $\mathbf{I}_H^h$ *is injective.*
- *The coarse grid matrix is given by* $\mathbf{A}_H := \mathbf{I}_h^H \mathbf{A}_h \mathbf{I}_H^h$.

Then we have:

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

- Define $\mathbf{Q}_{h,H} := \mathbf{I} - \mathbf{T}_{h,H} = \mathbf{I}_H^h \mathbf{A}_H^{-1} \mathbf{I}_h^H \mathbf{A}_h$. For all $\mathbf{z} \in \text{range}(\mathbf{I}_H^h)$ we have

$$\mathbf{Q}_{h,H}\mathbf{z} = \mathbf{Q}_{h,H}\mathbf{I}_H^h\mathbf{y} = \mathbf{I}_H^h \mathbf{A}_H^{-1} I_h^H \mathbf{A}_h \mathbf{I}_H^h \mathbf{y} = \mathbf{I}_H^h \mathbf{y} = \mathbf{z}.$$

  It follows (why?)

$$\text{null}(\mathbf{T}_{h,H}) = \text{range}(\mathbf{Q}_{h,H}) = \text{range}(\mathbf{I}_H^h).$$

- By Proposition 2, we know that $\text{range}(\mathbf{T}_{h,H})$ is orthogonal to

$$\text{null}(\mathbf{T}_{h,H}) = \text{range}(\mathbf{Q}_{h,H})$$

  with respect to $\langle \cdot, \cdot \rangle_{\mathbf{A}_h}$. Hence, we have an orthogonal decomposition

$$\mathbb{R}^{n_h} = \text{range}(\mathbf{T}_{h,H}) \oplus \text{null}(\mathbf{T}_{h,H}).$$

- Let $\mathbf{D}_h$ be the the diagonal part of $\mathbf{A}_h$. Note that we have

$$\|\mathbf{x}\|^2_{\mathbf{A}_h\mathbf{D}_h^{-1}\mathbf{A}_h} = \mathbf{x}^\top \mathbf{A}_h \mathbf{D}_h^{-1} \mathbf{A}_h \mathbf{x} = \|\mathbf{A}_h\mathbf{x}\|^2_{\mathbf{D}_h^{-1}}.$$

- We will say that the smoothing process $\mathbf{S}_h$ defined by

$$\mathbf{S}_h(\mathbf{x}) = \mathbf{R}_h\mathbf{x} + \mathbf{c}^h$$

has the *smoothing property* if there is a constant $\alpha > 0$ such that

$$\|\mathbf{R}_h\mathbf{v}^h\|^2_{\mathbf{A}_h} \leq \|\mathbf{v}^h\|^2_{\mathbf{A}_h} - \alpha\|\mathbf{A}_h\mathbf{v}^h\|^2_{\mathbf{D}_h^{-1}}, \qquad \mathbf{v}^h \in \mathbb{R}^{n_h}.$$

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

$$\min_{\mathbf{v}^H} \|\mathbf{v}^h - \mathbf{I}_H^h\mathbf{v}^H\|_{\mathbf{D}_h} \leq \beta\|\mathbf{v}^h\|_{\mathbf{A}_h\mathbf{D}_h^{-1}\mathbf{A}_h}, \qquad \mathbf{v}^h \in \mathbb{R}^{n_h}.$$

Theorem 3

Let the conditions of Proposition 2 be satisfied. Assume:

- *the smoothing process* $\mathbf{S}_h$ *has the smoothing property,*
- *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*

$$\|\mathbf{T}_h\|_{\mathbf{A}_h} \leq \sqrt{1 - \frac{\alpha}{\beta}}.$$

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

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

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

Output: Approximation to $\mathbf{A}_h^{-1}\mathbf{f}^h$.

1. `Presmooth` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}^h_{(0)})$
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 $\mathbf{A}_H\boldsymbol{\varepsilon}^H = \mathbf{r}^H$

   else

   $\boldsymbol{\varepsilon}^H :=$V-cycle$(\mathbf{A}_H, \mathbf{r}^H, \mathbf{0}, \ell_1, \ell_2, h_0)$

   end
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` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$

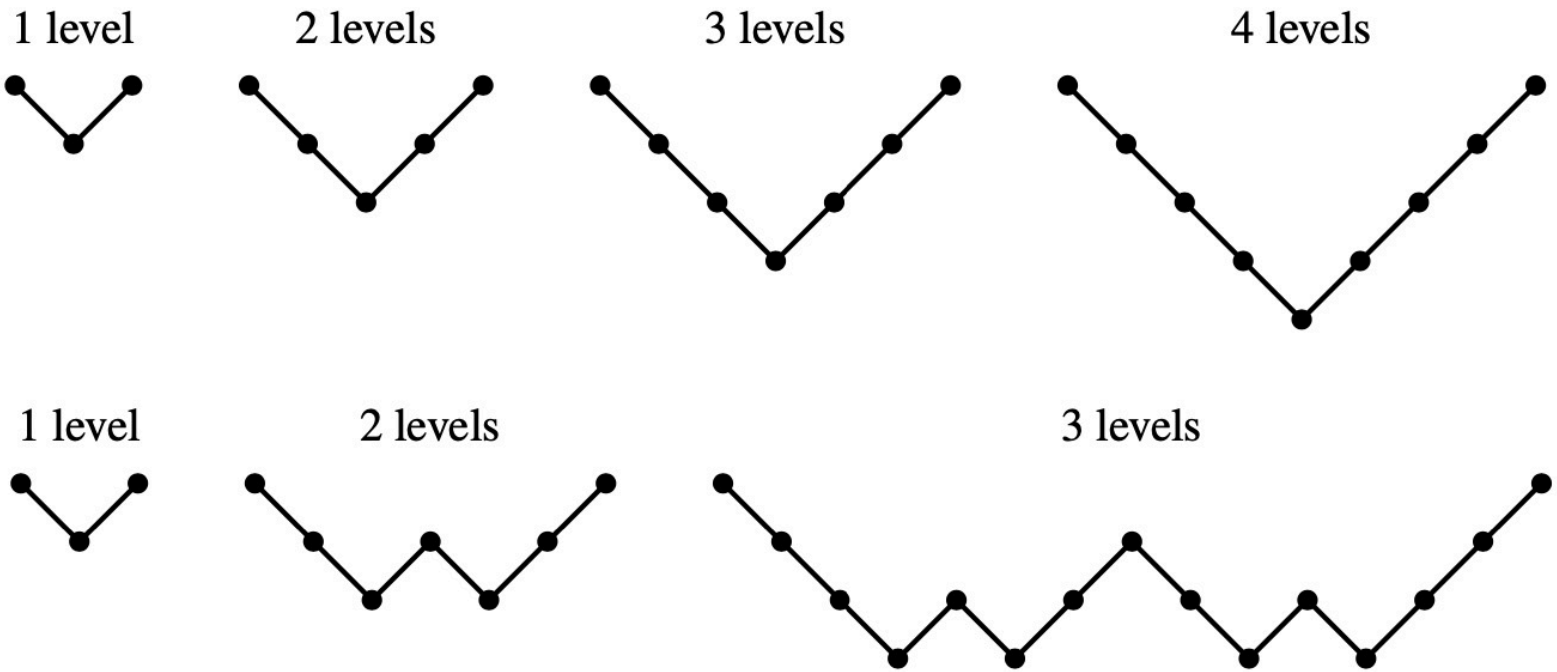
Algorithm: Multigrid for $\mathbf{A}_h\mathbf{u}^h = \mathbf{f}^h$, $\mathrm{MG}(\mathbf{A}_h, \mathbf{f}^h, \mathbf{u}^h_{(0)}, \ell_1, \ell_2, \ell, h_0)$

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

Output: Approximation to $\mathbf{A}_h^{-1}\mathbf{f}^h$.

1. `Presmooth` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_1}(\mathbf{u}^h_{(0)})$
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 $\mathbf{A}_H \boldsymbol{\varepsilon}^H = \mathbf{r}^H$
   else
   $\boldsymbol{\varepsilon}^H := \mathbf{0}$
   for $j = 1 : \ell$
   $\boldsymbol{\varepsilon}^H := \mathrm{MG}(\mathbf{A}_H, \mathbf{r}^H, \boldsymbol{\varepsilon}^H, \ell_1, \ell_2, \ell, h_0)$
   end
   end
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` : $\mathbf{u}^h := \mathbf{S}_h^{\ell_2}(\mathbf{u}^h)$

- Obviously, $\ell = 1$ leads to the V-cycle method. It is helpful to visualize the recursion in the following way, depending on the choice of ℓ and how many levels there are, meaning how many grids we use.



The recursion of multigrid with $\ell = 1$ (top) and $\ell = 2$ (bottom).

- 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 ℓ steps of the multigrid method on grid X_H:

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

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):
  - 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

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

  A Multigrid Tutorial

  Second Edition, SIAM, 2000

- Pieter Wesseling

  An Introduction to Multigrid Methods

  John Wiley & Sons, 1992