Notes on Multigrid Methods

Qinghai Zhang

March 22, 2018

Motivation of multigrids.

- The convergence rates of classical iterative method depend on the grid spacing, or problem size. In contrast, convergence rates of multigrid methods does not.
- The complexity is $O(n)$.

Textbook: Brigg, Henson, and McCormick 2000 A multigrid tutorial, SIAM, 2nd ed.
Encyclopedic website: `www.mgnet.org`

1 The model problem: 1D Possion equation.

On the unit 1D domain $x \in [0, 1]$, we numerically solve Poisson equation with homogeneous boundary condition

$$\Delta u = f, \qquad x(0) = x(1) = 0. \tag{1}$$

Discretize the domain with n cells and locate the knowns f_j and unknowns u_j at nodes $x_j = j/n = jh$, $j = 0, 1, \ldots, n$. We would like to approximate the second derivative of u using the discrete values at the nodes. Using Taylor expansion, we have

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{jh} = \frac{u_{j+1} + u_{j-1} - 2u_j}{h^2} + O(h^2). \tag{2}$$

Definition 1. *The* one-dimensional second-order discrete Laplacian *is a Toeplitz matrix $A \in \mathbb{R}^{(n-1)\times(n-1)}$ as*

$$a_{ij} = \begin{cases} 2, & i = j \\ -1, & i - j = \pm 1 \\ 0, & \textit{otherwise} \end{cases} \tag{3}$$

Then we are going to solve the linear system

$$A\mathbf{u} = \mathbf{f}, \tag{4}$$

where $f_j = h^2 f(x_j)$.

Proposition 2.

$$\frac{1}{h^2}(Au)_j - (\Delta u)|_{x_j} = O(h^2), \qquad \forall j = 1, \ldots, n-1. \tag{5}$$

Proposition 3. *The eigenvalues λ_k and eigenvectors $\mathbf{w}_k$ of A are*

$$\lambda_k(A) = 4\sin^2\frac{k\pi}{2n}, \tag{6}$$

$$w_{k,j} = \sin\frac{jk\pi}{n}, \tag{7}$$

where $j, k = 1, 2, \ldots, n-1$.

Proof. use the trigonmetric identity

$$\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}.$$

□

Remark 1. *The 2D counterpart of A is $A\otimes I + I\otimes A$.*

Note that this is not true for variable coefficient Poisson equation.

2 The residual equation

Definition 4. *For an approximate solution $\tilde{u}_j \approx u_j$, the* error *is $\mathbf{e} = \mathbf{u} - \tilde{\mathbf{u}}$, the* residual *is $\mathbf{r} = \mathbf{f} - A\tilde{\mathbf{u}}$. Then*

$$A\mathbf{e} = \mathbf{r} \tag{8}$$

holds and it is called the residual equation.

As one advantage, the residual equation lets us focus on homogenuous Dirichlet condition WLOG.

Question 1. *For inexact arithmetic, does a small residual imply a small error?*

Definition 5. *The condition number of a matrix A is $\mathrm{cond}(A) = \|A\|_2\|A^{-1}\|_2$. It indicates how well the residual measures the error.*

$$\|A\|_2 = \sup_{\mathbf{x}\neq 0}\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x}\neq 0}\sqrt{\frac{(A\mathbf{x},A\mathbf{x})}{(\mathbf{x},\mathbf{x})}} = \sup_{\mathbf{x}\neq 0}\sqrt{\frac{(\mathbf{x},A^TA\mathbf{x})}{(\mathbf{x},\mathbf{x})}} = \sqrt{\lambda_{\max}(A^TA)}$$

Since A is symmetric, $\|A\|_2 = \lambda_{\max}(A)$. $\|A^{-1}\|_2 = \lambda_{\max}(A^{-1}) = \lambda_{\min}^{-1}(A)$. To give you an idea about the magnitude of $\mathrm{cond}(A)$, for $n = 8$, cond(A)=32, for $n = 1024$, cond(A)=4.3e+5.

Theorem 6.

$$\frac{1}{\mathrm{cond}(A)}\frac{\|\mathbf{r}\|_2}{\|\mathbf{f}\|_2} \le \frac{\|\mathbf{e}\|_2}{\|\mathbf{u}\|_2} \le \mathrm{cond}(A)\frac{\|\mathbf{r}\|_2}{\|\mathbf{f}\|_2} \tag{9}$$

Proof. Use $\|A\mathbf{x}\| \le \|A\|\|\mathbf{x}\|$ and the equations $A\mathbf{e} = \mathbf{r}$, $A^{-1}\mathbf{f} = \mathbf{u}$. □

3 Fourier modes and aliasing

Hereafter Ω^h denote both the uniform grid with n intervals and the corresponding vector space.

Wavelength refers to the distance of one sinusoidal period.

Proposition 7. *The kth Fourier mode $w_{k,j} = \sin(x_jk\pi)$ has wavelength $L_w = \frac{2}{k}$.*

Proof. $\sin(x_jk\pi) = \sin(x_j + \frac{L}{2})k\pi$ implies $x_jk\pi = (x_j + \frac{L}{2})k\pi - \pi$. Hence $k = \frac{2}{L_w}$. □

The wavenumber k is the number of crests and troughs in the unit domain.

Question 2. *What is the range of representable wavenumbers on Ω^h?*

For $n = 8$, consider $k = 1, 2, 8$. $k_{\mathrm{rep}} \in [1, n)$. What happens to modes of $k > n$? E.g. the mode with $k = 3n/2$ is represented by $k = n/2$. Plot the case of $n = 4$.

Proposition 8. *On Ω^h, a Fourier mode $\mathbf{w}_k = \sin(x_jk\pi)$ with $n < k < 2n$ is actually represented as the mode $\mathbf{w}_{k'}$ where $k' = 2n - k$.*

Proof. $-\sin(x_jk\pi) = \sin(2j\pi - x_jk\pi) = \sin(x_j(2n-k)\pi) = \sin(x_jk'\pi) = w_{k'}$. □

Definition 9. *On Ω^h, the Fourier modes with wavenumbers $k \in [1, n/2)$ are called* low-frenquency *(LF) or* smooth *modes, those with $k \in [n/2, n-1)$* high-frenquency *(HF) or* oscillatory *modes.*

4 The spectral property of weighted Jacobi

The scalar fixed-point iteration converts the problem of finding a root of $f(x) = 0$ to the problem of finding a fixed point of $g(x) = x$ where $f(x) = c(g(x) - x)$ and $c \neq 0$.

Classical iterative methods split A as $A = M - N$ and convert (4) to a fixed point (FP) problem $\mathbf{u} = M^{-1}N\mathbf{u} + M^{-1}f$. Let $T = M^{-1}N$, $\mathbf{c} = M^{-1}\mathbf{f}$. Then fixed point iteration yields

$$\mathbf{u}^{(\ell+1)} = T\mathbf{u}^{(\ell)} + \mathbf{c}. \tag{10}$$

After ℓ iterations

$$\mathbf{e}^{(\ell)} = T^{\ell}\mathbf{e}^{(0)}. \tag{11}$$

Obviously, the FP iteration will converge iff the special radius $\rho(T) := |\lambda(T)|_{\max} < 1$.

Decompose A as $A = D + L + U$. Jacobi iteration has $M = D$, $N = -(L+U)$, $T = -D^{-1}(L+U)$, i.e.

$$t_{ij} = \begin{cases} \frac{1}{2}, & i - j = \pm 1 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Here $\rho(T) = 1 - O(h^2)$. As $h \to 0$, $\rho(T) \to 1$, and Jacobi converges slowly.

Consider a generalization of the Jacobi iteration.

Definition 10. *The* weighted Jacobi method *splits A as $A = D + L + U$ where D, L, U are diagonal, lower triangular, and upper triangular, respectively, and then performs fixed point iterations,*

$$\mathbf{u}^{*} = -D^{-1}(L+U)\mathbf{u}^{(\ell)} + D^{-1}\mathbf{f}, \tag{13a}$$

$$\mathbf{u}^{(\ell+1)} = (1-\omega)\mathbf{u}^{(\ell)} + \omega\mathbf{u}^{*}. \tag{13b}$$

Setting $\omega = 1$ yields Jacobi.

Proposition 11. *The weighted Jacobi has the iteration matrix*

$$T_{\omega} = (1-\omega)I - \omega D^{-1}(L+U) = I - \frac{\omega}{2}A, \tag{14}$$

whose eigenvectors are the same as those of A, with the corresponding eigenvalues as

$$\lambda_k(T_{\omega}) = 1 - 2\omega\sin^2\frac{k\pi}{2n}, \tag{15}$$

where $k = 1, 2, \ldots, n-1$.

See Fig. 2.7. For $n = 64$, $\omega \in [0, 1]$, $\rho(T_{\omega}) \geq 0.9986$. Not a great iteration method either. Why? Look more under the hood to consider how weighted Jacobi damps different modes. Write $\mathbf{e}^{(0)} = \sum_k c_k\mathbf{w}_k$, then

$$\mathbf{e}^{(\ell)} = T_{\omega}^{\ell}\mathbf{e}^{(0)} = \sum_k c_k\lambda_k^{\ell}(T_{\omega})\mathbf{w}_k. \tag{16}$$

No value of ω will reduce the smooth components of the error effectively.

$$\lambda_1(T_{\omega}) = 1 - 2\omega\sin^2\frac{\pi}{2n} \approx 1 - \frac{\omega\pi^2h^2}{2}. \tag{17}$$

Having accepted that no value of ω damps the smooth components satisfactorily, we ask what value of ω provides the best damping of the oscillatory modes.

Definition 12. *The* smoothing factor *μ is the maximal factor of damping for HF modes. An iterative method is said to have the* smoothing property *if μ is small and independent of the grid size.*

For weighted Jacobi, this optimization problem is

$$\mu = \min_{\omega\in(0,1]} \max_{k\in[n/2,n)} |\lambda_k(T_\omega)| . \tag{18}$$

$\lambda_k(T_\omega)$ is a monotonically decreasing function, and the minimum is therefore obtained by setting

$$\lambda_{n/2}(T_\omega) = -\lambda_n(T_\omega) \;\Rightarrow\; \omega = \frac{2}{3}. \tag{19}$$

Exercise:

$$\omega = \frac{2}{3} \;\Rightarrow\; |\lambda_k| \le \mu = \frac{1}{3} \tag{20}$$

See Figure 2.8 and 2.9. Regular Jacobi is only good for modes $16 \le k \le 48$. For $\omega = \frac{2}{3}$, the modes $16 \le k < 64$ are all damped out quickly.

5 Two-grid correction

5.1 The main idea and the linear operator

Proposition 13. *The kth LF mode on Ω^h is the kth mode on Ω^{2h}:*

$$w^h_{k,2j} = w^{2h}_{k,j}. \tag{21}$$

However, the LF modes $k \in [\frac{n}{4}, \frac{n}{2})$ of Ω^h will become HF modes on Ω^{2h}.

Proof.

$$w^h_{k,2j} = \sin\frac{2jk\pi}{n} = \sin\frac{jk\pi}{n/2} = w^{2h}_{k,j}, \tag{22}$$

where $k \in [1, n/2)$. Because of the smaller range of k on Ω^{2h}, the mode with $k \in [\frac{n}{4}, \frac{n}{2})$ are HF by definition since the highest wavenumber is $\frac{n}{2}$ on Ω^{2h}. $\square$

Definition 14. *The* restriction *operator $I_h^{2h} : \mathbb{R}^{n-1} \to \mathbb{R}^{n/2-1}$ maps a vector on the fine grid Ω^h to its counterpart on the coarse grid Ω^{2h}:*

$$I_h^{2h} v^h = v^{2h}. \tag{23}$$

A common restriction operator is the full-weighting *operator*

$$v_j^{2h} = \frac{1}{4}(v^h_{2j-1} + 2v^h_{2j} + v^h_{2j+1}), \tag{24}$$

where $j = 1, 2, \ldots, \frac{n}{2} - 1$.

Example 1. *For $n = 8$, the full-weighting operator is*

$$I_h^{2h} = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & \\ & & & & 1 & 2 & 1 \end{bmatrix}. \tag{25}$$

Definition 15. *The* prolongation or interpolation *operator $I_{2h}^h : \mathbb{R}^{n/2-1} \to \mathbb{R}^{n-1}$ maps a vector on the coarse grid Ω^{2h} to its counterpart on the fine grid Ω^h:*

$$I_{2h}^h v^{2h} = v^h. \tag{26}$$

A common prolongation is the linear interpolation *operator*

$$\begin{array}{rcl} v^h_{2j} &=& v^{2h}_j, \\ v^h_{2j+1} &=& \frac{1}{2}(v^{2h}_j + v^{2h}_{j+1}). \end{array} \tag{27}$$

Example 2. *For $n = 8$, the linear interpolation operator is*

$$I_{2h}^{h} = \frac{1}{2}\begin{bmatrix} 1 & & \\ 2 & & \\ 1 & 1 & \\ & 2 & \\ & 1 & 1 \\ & & 2 \\ & & 1 \end{bmatrix}. \tag{28}$$

Remark 2. *The weighted Jacobi with $\omega = \frac{2}{3}$ damps HF modes effectively. By Proposition 13, LF modes a fine grid may become HF modes on a coarse grid. This fact is exploited in multigrid methods on a series of successively coarsened grids to eliminate most HF modes.*

Definition 16. *For $A\mathbf{u} = \mathbf{f}$, the* two-grid correction scheme

$$v^h \leftarrow MG(v^h, f^h, \nu_1, \nu_2) \tag{29}$$

consists of the following steps:

1) Relax $A^h\mathbf{u}^h = f^h$ for ν_1 times on Ω^h with initial guess v^h: $v^h \leftarrow T_\omega^{\nu_1} v^h + \mathbf{c}'(f)$,

2) compute the fine-grid residual $r^h = f^h - A^h v^h$ and restrict it to the coarse grid by $r^{2h} = I_h^{2h} r^h$: $r^{2h} \leftarrow I_h^{2h}(f^h - A^h v^h)$,

3) solve $A^{2h}e^{2h} = r^{2h}$ on Ω^{2h}: $e^{2h} \leftarrow (A^{2h})^{-1} r^{2h}$,

4) interpolate the coarse-grid error to the fine grid by $e^h = I_{2h}^h e^{2h}$ and correct the fine-grid approximation: $v^h \leftarrow v^h + I_{2h}^h e^{2h}$,

5) Relax $A^h\mathbf{u}^h = f^h$ for ν_2 times on Ω^h with initial guess v^h: $v^h \leftarrow T_\omega^{\nu_2} v^h + \mathbf{c}'(f)$.

Remark 3. *The* type *of boundary conditions is incorporated in the matrix A^h while the* value *of boundary conditions is incorporated in the vector $\mathbf{c}'(f)$. In step 3), we solve for the* exact *value because later in the recursive version we will coarsen the grid to an extend that its number of cells is a small integer such as 4 or 8.*

Proposition 17. *Let TG denote the iteration matrix of the two-grid correction scheme acting on the error vector. Then*

$$TG = T_\omega^{\nu_2}\left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] T_\omega^{\nu_1}. \tag{30}$$

Proof. By definition, the two-grid correction scheme with $\nu_2 = 0$ replaces the initial guess with

$$v^h \leftarrow T_\omega^{\nu_1} v^h + \mathbf{c}'(f) + I_{2h}^h (A^{2h})^{-1} I_h^{2h}\left[f^h - A^h(T_\omega^{\nu_1} v^h + \mathbf{c}'(f))\right], \tag{31}$$

which also holds for the exact solution u^h

$$u^h \leftarrow T_\omega^{\nu_1} u^h + \mathbf{c}'(f) + I_{2h}^h (A^{2h})^{-1} I_h^{2h}\left[f^h - A^h(T_\omega^{\nu_1} u^h + \mathbf{c}'(f))\right]$$

Subtracting the two equations yields

$$e^h \leftarrow T_\omega^{\nu_1} e^h + I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h T_\omega^{\nu_1} e^h.$$

Similar arguments applied to step 5) yield (30). □

5.2 The spectral picture

Our objective is to show that $\rho(TG) \approx 0.1$ for $\nu_1 = 2, \nu_2 = 0$. For this purpose, we need to examine the intergrid transfer operators.

Definition 18. *$\mathbf{w}_k^h$ ($k \in [1, n/2)$) and $\mathbf{w}_{k'}^h$ ($k' = n - k$) are called* complementary modes *on Ω^h.*

Proposition 19. *For a pair of complementary modes on Ω^h, we have*

$$w_{k',j}^h = (-1)^{j+1} w_{k,j}^h \tag{32}$$

Proof. $w_{k',j}^h = \sin\frac{(n-k)j\pi}{n} = \sin\left(j\pi - \frac{kj\pi}{n}\right) = (-1)^{j+1} w_{k,j}^h$. ☐

Lemma 20. *The action of the full-weighting operator on a pair of complementary modes is*

$$I_h^{2h}\mathbf{w}_k^h = \cos^2\frac{k\pi}{2n}\mathbf{w}_k^{2h} := c_k\mathbf{w}_k^{2h}, \tag{33a}$$

$$I_h^{2h}\mathbf{w}_{k'}^h = -\sin^2\frac{k\pi}{2n}\mathbf{w}_k^{2h} := -s_k\mathbf{w}_k^{2h}, \tag{33b}$$

where $k \in [1, n/2)$, $k' = n - k$. In addition, $I_h^{2h}\mathbf{w}_{n/2}^h = 0$.

Proof. For the smooth mode,

$$\left(I_h^{2h}\mathbf{w}_k^h\right)_j = \frac{1}{4}\sin\frac{(j-1)k\pi}{n} + \frac{1}{2}\sin\frac{jk\pi}{n} + \frac{1}{4}\sin\frac{(j+1)k\pi}{n} = \frac{1}{2}\left(1+\cos\frac{k\pi}{n}\right)\sin\frac{jk\pi}{n} = \cos^2\frac{k\pi}{2n}w_{k,j}^{2h},$$

where the last step uses Proposition 13. As for the HF mode, follow the same procedure, but replace k with $n-k$, use Proposition 8 for aliasing, and notice that j is even. ☐

The full-weighting operator thus maps a pair of complementary modes to a multiple of the smooth mode, which might be an HF mode on the coarse grid.

Lemma 21. *The action of the interpolation operator on Ω^{2h} is*

$$I_{2h}^h\mathbf{w}_k^{2h} = c_k\mathbf{w}_k^h - s_k\mathbf{w}_{k'}^h, \tag{34}$$

where $k' = n - k$.

Proof. Proposition 19 and trignometric identities yield

$$c_k\mathbf{w}_k^h - s_k\mathbf{w}_{k'}^h = \left(\cos^2\frac{k\pi}{2n} + (-1)^j\sin^2\frac{k\pi}{2n}\right)\mathbf{w}_k^h = \begin{cases} \mathbf{w}_k^h, & \text{for even } j\text{'s}, \\ \cos\frac{k\pi}{n}\mathbf{w}_k^h, & \text{for odd } j\text{'s}. \end{cases}$$

On the other hand, by Definition 15, we have

$$\left(I_{2h}^h\mathbf{w}_k^{2h}\right)_j = \begin{cases} w_{k,j}^h, & j \text{ is even}, \\ \frac{1}{2}\sin\frac{k\pi(j-1)/2}{n/2} + \frac{1}{2}\sin\frac{k\pi(j+1)/2}{n/2} = \cos\frac{k\pi}{n}w_{k,j}^h, & j \text{ is odd}. \end{cases}$$ ☐

Remark 4. *By Lemma 21, the range of the interpolation operator contains both smooth and oscillatory modes. In other words, it excites oscillatory modes on the fine grid. However, if $k \ll n$, the amplitudes of these HF modes are small: $s_k \sim O(\frac{k^2}{n^2})$.*

Theorem 22. *The two-grid correction operator is invariant on the subspace $W_k^h = \mathrm{span}\{\mathbf{w}_k^h, \mathbf{w}_{k'}^h\}$.*

$$TG\mathbf{w}_k = \lambda_k^{\nu_1+\nu_2}s_k\mathbf{w}_k + \lambda_k^{\nu_1}\lambda_{k'}^{\nu_2}s_k\mathbf{w}_{k'} \tag{35a}$$

$$TG\mathbf{w}_{k'} = \lambda_{k'}^{\nu_1}\lambda_k^{\nu_2}c_k\mathbf{w}_k + \lambda_{k'}^{\nu_1+\nu_2}c_k\mathbf{w}_{k'}, \tag{35b}$$

where λ_k is the eigenvalue of T_ω.

Proof. Consider first the case of $\nu_1 = \nu_2 = 0$.

$$
\begin{aligned}
&A^h \mathbf{w}_k^h = 4 s_k \mathbf{w}_k^h && \text{(36a)}\\
\Rightarrow & I_h^{2h} A^h \mathbf{w}_k^h = 16 c_k s_k \mathbf{w}_k^{2h} && \text{(36b)}\\
\Rightarrow & (A^{2h})^{-1} I_h^{2h} A^h \mathbf{w}_k^h = \frac{16 c_k s_k}{4\sin^2 \frac{k\pi}{n/2}} \mathbf{w}_k^{2h} = \mathbf{w}_k^{2h} && \text{(36c)}\\
\Rightarrow & - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h \mathbf{w}_k^h = -c_k \mathbf{w}_k^h + s_k \mathbf{w}_{k'}^h && \text{(36d)}\\
\Rightarrow & \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h \right] \mathbf{w}_k^h = s_k \mathbf{w}_k^h + s_k \mathbf{w}_{k'}^h, && \text{(36e)}
\end{aligned}
$$

where the additional factor of 4 in (36b) comes from the fact that the residual is scaled by h^2 and the trigonmetric identity $\sin(2\theta) = 2\sin\theta\cos\theta$ is applied in (36c). Similarly,

$$
\begin{aligned}
&A^h \mathbf{w}_{k'}^h = 4 s_{k'} \mathbf{w}_{k'}^h = 4 c_k \mathbf{w}_{k'}^h && \text{(37a)}\\
\Rightarrow & I_h^{2h} A^h \mathbf{w}_{k'}^h = -16 c_k s_k \mathbf{w}_k^{2h} && \text{(37b)}\\
\Rightarrow & (A^{2h})^{-1} I_h^{2h} A^h \mathbf{w}_{k'}^h = -\frac{16 c_k s_k}{4\sin^2 \frac{k\pi}{n/2}} \mathbf{w}_k^{2h} = -\mathbf{w}_k^{2h} && \text{(37c)}\\
\Rightarrow & - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h \mathbf{w}_{k'}^h = c_k \mathbf{w}_k^h - s_k \mathbf{w}_{k'}^h && \text{(37d)}\\
\Rightarrow & \left(I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h \right) \mathbf{w}_{k'}^h = c_k \mathbf{w}_k^h + c_k \mathbf{w}_{k'}^h. && \text{(37e)}
\end{aligned}
$$

Note that in the first equation we have used $c_k = s_{k'}$.

Adding pre-smoothing incurs a scaling of $\lambda_k^{\nu_1}$ for (36e) and $\lambda_{k'}^{\nu_1}$ for (37e). In contrast, adding post-smoothing incurs a scaling of $\lambda_k^{\nu_2}$ for $\mathbf{w}_k^h$ and a scaling of $\lambda_{k'}^{\nu_2}$ for $\mathbf{w}_{k'}^h$ in both (36e) and (37e). Hence (35) holds. □

Remark 5. *(35) can be rewritten as*

$$
TG \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k'} \end{bmatrix} = \begin{bmatrix} \lambda_k^{\nu_1+\nu_2} s_k & \lambda_k^{\nu_1}\lambda_{k'}^{\nu_2} s_k \\ \lambda_{k'}^{\nu_1}\lambda_k^{\nu_2} c_k & \lambda_{k'}^{\nu_1+\nu_2} c_k \end{bmatrix} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k'} \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k'} \end{bmatrix}. \tag{38}
$$

For $k \ll n$, although $\lambda_k^{\nu_1+\nu_2} \approx 1$, $s_k \sim \frac{k^2}{n^2}$, hence $c_1 \ll 1$. Also, $\lambda_{k'}^{\nu_1} \ll 1$, hence $c_2, c_3, c_4 \ll 1$. See Figure 1 for four examples.

5.3 The algebraic picture

Lemma 23. *The full-weighting operator and the linear-interpolation operator satisfy the* variational properties

$$
I_{2h}^h = c(I_h^{2h})^T, \ c \in \mathbb{R}. \tag{39a}
$$

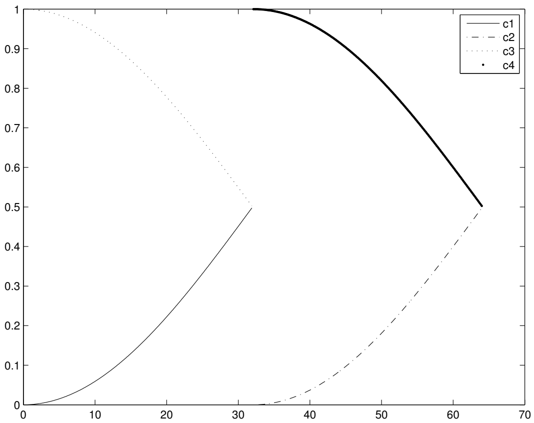
$$
I_h^{2h} A^h I_{2h}^h = A^{2h}. \tag{39b}
$$

(39b) is also called the Galerkin condition*.*

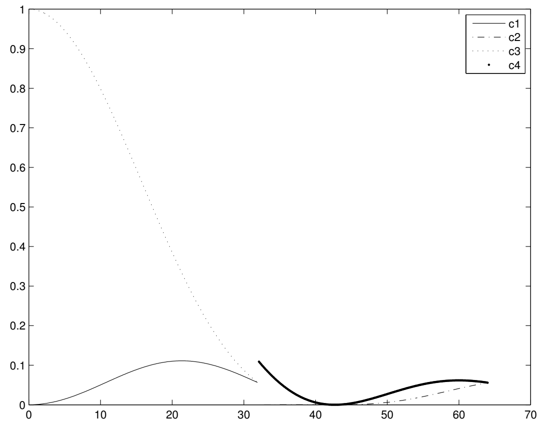
Proposition 24. *A basis for the range of the interpolation operator $\mathcal{R}(I_{2h}^h)$ is given by its columns, hence dim $\mathcal{R}(I_{2h}^h) = \frac{n}{2} - 1$. Its null space $\mathcal{N}(I_{2h}^h) = \{\mathbf{0}\}$.*

Proof. $\mathcal{R}(I_{2h}^h) = \{I_{2h}^h v^{2h} : v^{2h} \in \Omega^{2h}\}$. The maximum dimension of $\mathcal{R}(I_{2h}^h)$ is thus $\frac{n}{2} - 1$. Any v^{2h} can be expressed as $v^{2h} = \sum v_j^{2h} \mathbf{e}_j^{2h}$. It is obvious that the columns of I_{2h}^h are linearly independent. □

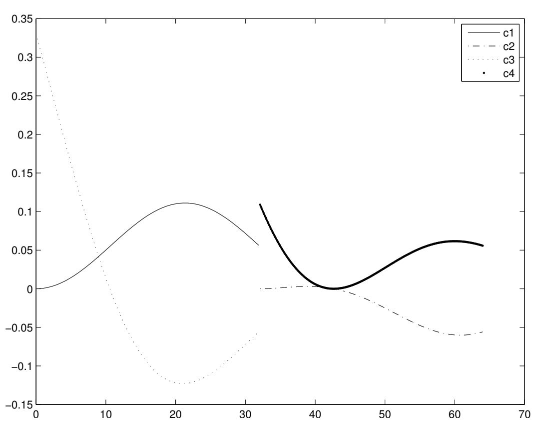
Definition 25. *For a matrix $A \in \mathbb{R}^{m\times n} : \mathbb{R}^n \to \mathbb{R}^m$, its* column space *consists of all linear combinations of its columns, while its* row space *is the column space of A^T. The* null space *is the set of vectors $\mathcal{N}(A) = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$. The* left null space *is the null space of A^T.*



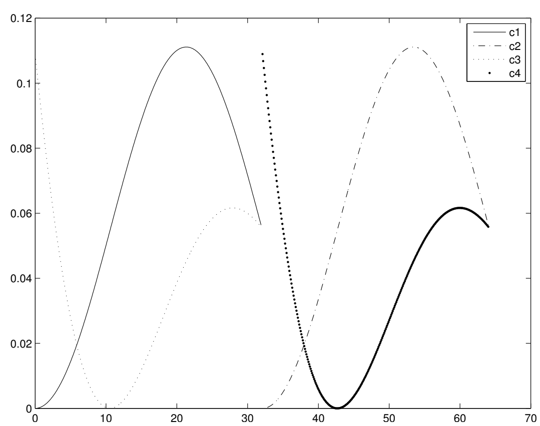
(a) $\nu_1 = 0$, $\nu_2 = 0$



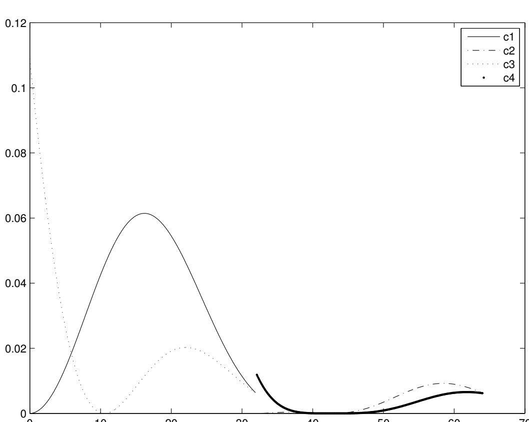
(b) $\nu_1 = 0$, $\nu_2 = 2$



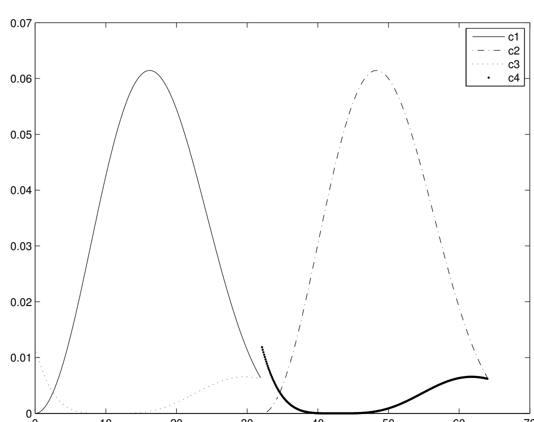
(c) $\nu_1 = 1$, $\nu_2 = 1$



(d) $\nu_1 = 2$, $\nu_2 = 0$



(e) $\nu_1 = 2$, $\nu_2 = 2$



(f) $\nu_1 = 4$, $\nu_2 = 0$

Figure 1: The damping coefficients of two-grid correction with weighted Jacobi for $n = 64$. The x-axis is k.

Theorem 26 (The counting theorem or the fundamental theorem of linear maps)**.** *Suppose the vector space $\mathcal{V}$ is finite-dimensional and $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Then the range of T is finite-dimensional and*

$$\dim \mathcal{V} = \dim \mathcal{N}(T) + \dim \mathcal{R}(T).$$

Theorem 27 (Fundamental theorem of linear algebra)**.** *For a matrix $A \in \mathbb{R}^{m\times n} : \mathbb{R}^n \to \mathbb{R}^m$, its column space and row space both have dimension r. The null spaces have dimensions $n-r$ and $m-r$. In addition, we have*

$$\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T), \tag{40a}$$
$$\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A), \tag{40b}$$

where $\mathcal{R}(A) \perp \mathcal{N}(A^T)$ and $\mathcal{R}(A^T) \perp \mathcal{N}(A)$.

Proof. $\mathbf{x} \in \mathcal{N}(A)$ implies $\mathbf{x} \in \mathbb{R}^n$ and $A\mathbf{x} = \mathbf{0}$. The latter expands to

$$\begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \cdots \\ \mathbf{a}_m^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix},$$

which implies $\forall j = 1, 2, \ldots, m, \qquad \mathbf{a}_j \perp \mathbf{x}$. Hence $\mathbf{x}$ is orthorgonal to each basis vector of $\mathcal{R}(A^T)$. □

Remark 6. *Each $\mathbf{x} \in \mathbb{R}^n$ can be split into a row space component $\mathbf{x}_r$ and a null space component $\mathbf{x}_n$. Then $A\mathbf{x} = A\mathbf{x}_r \in \mathcal{R}(A)$.* Every vector goes to the column space! *Furthermore,* every vector in the column space comes from one and only one vector in the row space.

Corollary 28. *For the full-weighting operator,*

$$\dim \mathcal{R}(I_h^{2h}) = \frac{n}{2} - 1, \qquad \dim \mathcal{N}(I_h^{2h}) = \frac{n}{2}. \tag{41}$$

Proof. See Figure 5.7 on page 85. The rest of the proof follows from (39). □

Remark 7. *If A has rank r, from the singular value decomposition $A = U\Sigma V^T$, we have*

$$\mathcal{R}(A) = \operatorname{span}\{U_1, U_2, \ldots, U_r\}, \tag{42}$$
$$\mathcal{N}(A) = \operatorname{span}\{V_{n-r+1}, V_{n-r+2}, \ldots, V_n\}, \tag{43}$$
$$\mathcal{R}(A^T) = \operatorname{span}\{V_1, V_2, \ldots, V_r\}, \tag{44}$$
$$\mathcal{N}(A^T) = \operatorname{span}\{U_{m-r+1}, U_{m-r+2}, \ldots, U_m\}. \tag{45}$$

This is closely related to Theorem 27.

Proposition 29. *A basis for the null space of the full-weighting operator is given by*

$$\mathcal{N}(I_h^{2h}) = \operatorname{span}\{A^h \mathbf{e}_j^h : j \text{ is odd}\}, \tag{46}$$

where $\mathbf{e}_j^h$ is the jth unit vector on Ω^h.

Proof. Consider $I_h^{2h}A^h$. The jth row of I_h^{2h} has $2(j-1)$ leading zeros and the next three nonzero entries are $1/4, 1/2, 1/4$. Since A^h has bandwidth of 3, it suffices to only consider five columns of A^h for potentially non-zero dot-product $\sum_i (I_h^{2h})_{ji}(A^h)_{ik}$. For $2j \pm 1$, these dot products are zero; for $2j$, the dot product is $1/2$; for $2j \pm 2$, the dot product is $-1/4$;

Hence for any odd j, we have $I_h^{2h}A^h\mathbf{e}_j^h = \mathbf{0}$. □

The above proposition states that the basis vector of $\mathcal{N}(I_h^{2h})$ are of the form

$$(0, 0, \dots, -1, 2, -1, \dots, 0, 0)^T;$$

see Figure 5.4 on page 81. Hence $\mathcal{N}(I_h^{2h})$ consists of both smooth and oscillatory modes.

Theorem 30. *The null space of the two-grid correction operator is the range of interpolation:*

$$\mathcal{N}(TG) = \mathcal{R}(I_{2h}^h). \tag{47}$$

Proof. If $\mathbf{s}^h \in \mathcal{R}(I_{2h}^h)$, then $\mathbf{s}^h = I_{2h}^h \mathbf{q}^{2h}$.

$$TG\mathbf{s}^h = \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] I_{2h}^h \mathbf{q}^{2h} = \mathbf{0},$$

where the last step comes from (39b). Hence $\mathcal{R}(I_{2h}^h) \subseteq \mathcal{N}(TG)$.

By Proposition 29, $\mathbf{t}^h \in \mathcal{N}(I_h^{2h} A^h)$ implies that $\mathbf{t}^h = \sum_{j \text{ is odd}} t_j \mathbf{e}_j$. Consequently,

$$TG\mathbf{t}^h = \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] \mathbf{t}^h = \mathbf{t}^h,$$

i.e., TG is the identity operator when acting on $\mathcal{N}(I_h^{2h} A^h)$. Hence the dimension of $\mathcal{N}(TG)$ is no greater than the dimension of $\mathcal{R}(I_h^{2h} A^h)$, which is the same as $\dim \mathcal{R}(I_{2h}^h)$ since A^h is a bijection with full rank on $\mathbb{R}^{n-1}$. This implies that $\dim\mathcal{N}(TG) \le \dim\mathcal{R}(I_{2h}^h)$, which completes the proof. □

Now that we have both the spectral decomposition $\Omega^h = L \oplus H$ and the subspace decomposition $\Omega^h = \mathcal{R}(I_{2h}^h) \oplus \mathcal{N}(I_h^{2h} A^h)$, the combination of relaxation with TG correction is equivalent to projecting the initial error vector to the L axis and then to the $\mathcal{N}$ axis. Repeating this process reduces the error vector to the origin; see Figure 5.8-Figure 5.11 for an illustration.

6 Multigrid cycles

Definition 31. *The* V-cycle scheme *is an algorithm*

$$\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h, \nu_1, \nu_2) \tag{48}$$

with the following steps.

1) relax ν_1 times on $A^h \mathbf{u}^h = \mathbf{f}^h$ with a given initial guess $\mathbf{v}^h$,

2) if Ω^h is the coarsest grid, go to step 4), otherwise

$$\begin{aligned} \mathbf{f}^{2h} &\leftarrow I_h^{2h}(\mathbf{f}^h - A\mathbf{v}^h), \\ \mathbf{v}^{2h} &\leftarrow \mathbf{0}, \\ \mathbf{v}^{2h} &\leftarrow V^h(\mathbf{v}^{2h}, \mathbf{f}^{2h}). \end{aligned}$$

3) interpolate error back and correct the solution: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h \mathbf{v}^{2h}$.

4) relax ν_2 times on $A^h \mathbf{u}^h = \mathbf{f}^h$ with the initial guess $\mathbf{v}^h$.

Definition 32. *The* Full Multigrid V-cycle *is an algorithm*

$$\mathbf{v}^h \leftarrow FMG^h(\mathbf{f}^h, \nu_1, \nu_2) \tag{49}$$

with the following steps.

1) If Ω^h is the coarsest grid, set $\mathbf{v}^h \leftarrow \mathbf{0}$ and go to step 3), otherwise

$$\mathbf{f}^{2h} \leftarrow I_h^{2h}\mathbf{f}^{\mathbf{h}},$$
$$\mathbf{v}^{2h} \leftarrow FMG^{2h}(\mathbf{f}^{2h}, \nu_1, \nu_2).$$

2) correct $\mathbf{v}^h \leftarrow I_{2h}^h \mathbf{v}^{2h}$,

3) perform a V-cycle with initial guess $\mathbf{v}^h$: $\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h, \nu_1, \nu_2)$.

See Figure 3.6 for the above two methods. Note that in Figure 3.6(c) the initial descending to the coarsest grid is missing.