Notes on Multigrid Methods

Qinghai Zhang

April 23, 2018

Motivation of multigrids.

• The convergence rates of classical iterative methods depend on the grid spacing, or problem size. In contrast, convergence rates of multigrid methods do not.

• The complexity is $O(n)$.


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

$$\Delta u = f, \quad x(0) = x(1) = 0. \quad (1)$$

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

$$\frac{\partial^2 u}{\partial x^2} \bigg|_{x_j} = \frac{u_{j+1} + u_{j-1} - 2u_j}{h^2} + O(h^2). \quad (2)$$

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

$$a_{ij} = \begin{cases} 2, & i = j \\ -1, & i - j = \pm 1 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Then we are going to solve the linear system

$$Au = f, \quad (4)$$

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

**Proposition 2.**

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

**Proposition 3.** The eigenvalues $\lambda_k$ and eigenvectors $w_k$ of $A$ are

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

$$w_{k,j} = \sin \frac{jk\pi}{n}, \quad (7)$$

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


Theorem 6. The magnitude of cond(A) is symmetric, the residual measures the error.

Proof. use the trigonometric 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 ⊗ I + I ⊗ 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 e = u - \( \tilde{u} \), the residual is r = f - A\( \tilde{u} \).

Then

\[ Ae = r \tag{8} \]

holds and it is called the residual equation.

As one advantage, the residual equation lets us focus on homogenous 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 cond(A) = \( \| A \|_2 \| A^{-1} \|_2 \). It indicates how well the residual measures the error.

\[ \| A \|_2 = \sup_{x \neq 0} \frac{\| Ax \|_2}{\| x \|_2} = \sup_{x \neq 0} \sqrt{\frac{(Ax, Ax)}{(x, x)}} = \sup_{x \neq 0} \sqrt{\frac{(x, A^T Ax)}{(x, x)}} = \sqrt{\lambda_{\text{max}}(A^T A)} \]

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

Theorem 6.

\[ \frac{1}{\text{cond}(A)} \| r \|_2 \leq \| e \|_2 \leq \text{cond}(A) \| r \|_2 \tag{9} \]

Proof. Use \( \| Ax \| \leq \| A \| \| x \| \) and the equations \( Ae = r \), \( A^{-1}f = u \).

3 Fourier modes and aliasing

Hereafter \( \Omega^h \) 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 \( w_{k,j} = \sin(x_j k \pi) \) has wavelength \( L_w = \frac{2}{k} \).

Proof. \( \sin(x_j k \pi) = \sin(x_j + \frac{1}{2}) k \pi \) implies \( x_j k \pi = (x_j + \frac{1}{2}) k \pi - \pi \). 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 \( \Omega^h \)?

For \( n = 8 \), consider \( k = 1, 2, 8 \). \( k_{\text{rep}} \in [1,n] \). 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 \( \Omega^h \), a Fourier mode \( w_k = \sin(x_j k \pi) \) with \( n < k < 2n \) is actually represented as the mode \( w_{k'} \) where \( k' = 2n - k \).

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

Definition 9. On \( \Omega^h \), the Fourier modes with wavenumbers \( k \in [1,n/2] \) are called low-frequency (LF) or smooth modes, those with \( k \in [n/2,n-1] \) high-frequency (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 \( u = M^{-1}Nu + M^{-1}f \). Let \( T = M^{-1}N, \ c = M^{-1}f \). Then fixed point iteration yields

\[
\mathbf{u}^{(\ell+1)} = T\mathbf{u}^{(\ell)} + \mathbf{c}.
\]

(10)

After \( \ell \) iterations

\[
\mathbf{e}^{(\ell)} = T^\ell\mathbf{e}^{(0)}.
\]

(11)

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

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

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

(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,

\[
\mathbf{u}^* = -D^{-1}(L + U)\mathbf{u}^{(\ell)} + D^{-1}\mathbf{f},
\]

(13a)

\[
\mathbf{u}^{(\ell+1)} = (1 - \omega)\mathbf{u}^{(\ell)} + \omega\mathbf{u}^*.
\]

(13b)

Setting \( \omega = 1 \) yields Jacobi.

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

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

(14)

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

\[
\lambda_k(T_\omega) = 1 - 2\omega \sin^2 \frac{k\pi}{2n},
\]

(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

\[
\mathbf{e}^{(\ell)} = T_\omega^{\ell}\mathbf{e}^{(0)} = \sum_k c_k \lambda_k^{\ell}(T_\omega)\mathbf{w}_k.
\]

(16)

No value of \( \omega \) will reduce the smooth components of the error effectively.

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

(17)

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

**Definition 12.** The smoothing factor \( \mu \) is the maximal factor of damping for HF modes. An iterative method is said to have the smoothing property if \( \mu \) is small and independent of the grid size.
For weighted Jacobi, this optimization problem is
\[
\mu = \min_{\omega \in [0,1]} \max_{k \in [n/2,n)} |\lambda_k(T_\omega)|.
\] (18)
\[\lambda_k(T_\omega)\] is a monotonically decreasing function, and the minimum is therefore obtained by setting
\[
\lambda_{n/2}(T_\omega) = -\lambda_n(T_\omega) \Rightarrow \omega = \frac{2}{3}.
\] (19)

Exercise:
\[\omega = \frac{2}{3} \Rightarrow |\lambda_k| \leq \mu = \frac{1}{3} \] (20)

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

5 Two-grid correction

5.1 The main idea and the linear operator

Proposition 13. The \(k\)th LF mode on \(\Omega^h\) is the \(k\)th mode on \(\Omega^{2h}\):
\[
w^h_{k,j} = w^{2h}_{k,j}.
\] (21)
However, the LF modes \(k \in \left[\frac{n}{4}, \frac{n}{2}\right)\) of \(\Omega^h\) will become HF modes on \(\Omega^{2h}\).

Proof.
\[
w^h_{k,j} = \sin \frac{2jk\pi}{n} = \sin \frac{jk\pi}{n/2} = w^{2h}_{k,j},
\] (22)
where \(k \in [1, n/2)\). Because of the smaller range of \(k\) on \(\Omega^{2h}\), the mode with \(k \in \left[\frac{n}{4}, \frac{n}{2}\right)\) are HF by definition since the highest wavenumber is \(\frac{n}{2}\) on \(\Omega^{2h}\).

Definition 14. The restriction operator \(I^h_{2h} : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n/2-1}\) maps a vector on the fine grid \(\Omega^h\) to its counterpart on the coarse grid \(\Omega^{2h}\):
\[
I^h_{2h} v^h = v^{2h}.
\] (23)
A common restriction operator is the full-weighting operator
\[
v^{2h}_j = \frac{1}{4} (v^h_{2j-1} + 2v^h_{2j} + v^h_{2j+1}),
\] (24)
where \(j = 1, 2, \ldots, \frac{n}{2} - 1\).

Example 1. For \(n = 8\), the full-weighting operator is
\[
I^h_{2h} = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 & 1 & 2 & 1 & 1 & 2 \end{bmatrix}.
\] (25)

Definition 15. The prolongation or interpolation operator \(I^h_{2h} : \mathbb{R}^{n/2-1} \rightarrow \mathbb{R}^{n-1}\) maps a vector on the coarse grid \(\Omega^{2h}\) to its counterpart on the fine grid \(\Omega^h\):
\[
I^h_{2h} v^{2h} = v^h.
\] (26)
A common prolongation is the linear interpolation operator
\[
v^h_{2j} = v^{2h}_j,
v^h_{2j+1} = \frac{1}{2} (v^{2h}_j + v^{2h}_{j+1}).
\] (27)
Example 2. For $n = 8$, the linear interpolation operator is

$$I^h_{2h} = \frac{1}{2} \begin{bmatrix} 1 & 2 & 1 & 1 & 2 & 1 & 2 & 1 \end{bmatrix}.$$  \hspace{1cm} (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 $Au = f$, the two-grid correction scheme

$$v^h \leftarrow MG(v^h, f^h, \nu_1, \nu_2)$$

consists of the following steps:

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

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

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

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

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

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 $c'(f)$. 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

$$TG = T_{\nu_2}^h [I - I_{2h}^h (A^{2h})^{-1} I_{2h}^h A^h] T_{\nu_1}^h.$$  \hspace{1cm} (30)

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

$$v^h \leftarrow T_{\nu_1}^h v^h + c'(f) + I_{2h}^h (A^{2h})^{-1} I_{2h}^h [f^h - A^h (T_{\nu_1}^h v^h + c'(f))],$$  \hspace{1cm} (31)

which also holds for the exact solution $u^h$

$$u^h \leftarrow T_{\nu_1}^h u^h + c'(f) + I_{2h}^h (A^{2h})^{-1} I_{2h}^h [f^h - A^h (T_{\nu_1}^h u^h + c'(f))].$$

Subtracting the two equations yields

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

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.** $w^h_k \ (k \in [1, n/2])$ and $w^h_{k'} \ (k' = n - k)$ are called complementary modes on $\Omega^h$.

**Proposition 19.** For a pair of complementary modes on $\Omega^h$, we have
\[ w^h_{k'j} = (-1)^{j+1}w^h_{kj} \]

*Proof.* $w^h_{k'j} = \sin \left( \frac{(n-k)j\pi}{n} \right) = \sin \left( j\pi - \frac{k\pi}{n} \right) = (-1)^{j+1}w^h_{kj}$. \hfill $\square$

**Lemma 20.** The action of the full-weighting operator on a pair of complementary modes is
\[ I^{2h}_h w^h_k = \cos^2 \frac{k\pi}{2n} w^h_k := c_k w^h_k, \]
\[ I^{2h}_h w^h_{k'} = -\sin^2 \frac{k\pi}{2n} w^h_k := -s_k w^h_k, \]
where $k \in [1, n/2], k' = n - k$. In addition, $I^{2h}_h w^h_{n/2} = 0$.

*Proof.* For the smooth mode,
\[ (I^{2h}_h w^h_k)_j = \frac{1}{4} \sin \left( \frac{j-1}{n} \right) k \pi + \frac{1}{2} \sin \frac{j}{n} k \pi + \frac{1}{4} \sin \left( \frac{j+1}{n} \right) k \pi = \frac{1}{2} \left( 1 + \cos \frac{k\pi}{n} \right) \sin \frac{j\pi}{n} = \cos^2 \frac{k\pi}{2n} w^h_{k,j}, \]
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. \hfill $\square$

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 $\Omega^{2h}$ is
\[ I^{2h}_h w^h_k = c_k w^h_k - s_k w^h_{k'}, \]
where $k' = n - k$.

*Proof.* Proposition 19 and trigonometric identities yield
\[ c_k w^h_k - s_k w^h_{k'} = \left( \cos^2 \frac{k\pi}{2n} + (-1)^j \sin^2 \frac{k\pi}{2n} \right) w^h_k = \begin{cases} w^h_k, & \text{for even } j, \\ \cos^2 \frac{k\pi}{2n} w^h_{k,j}, & \text{for odd } j. \end{cases} \]

On the other hand, by Definition 15, we have
\[ (I^{2h}_h w^h_k)_j = \begin{cases} \frac{1}{2} \sin \frac{k\pi(1/j+1)}{n/2} + \frac{1}{2} \sin \frac{k\pi(j+1)/2}{n/2} = \cos \frac{k\pi}{n} w^h_{k,j}, & \text{if } \text{even}, \\ \frac{1}{2} \sin \frac{k\pi(1/j+1)}{n/2} - \frac{1}{2} \sin \frac{k\pi(j+1)/2}{n/2} = -\cos \frac{k\pi}{n} w^h_{k,j}, & \text{if } \text{odd}. \end{cases} \]

**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 $W^h_k = \text{span}\{w^h_k, w^h_{k'}\}$.
\[ TG w_k = \lambda_k^{\nu_2} s_k w_k + \lambda_k^{\nu_2} s_k w_{k'}, \]
\[ TG w_{k'} = \lambda_k^{\nu_2} s_k w_k + \lambda_k^{\nu_2} c_k w_{k'}, \]
where $\lambda_k$ is the eigenvalue of $T\omega$. 


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

\begin{align}
A^h w_k^h &= 4s_k w_k^h \quad (36a) \\
\Rightarrow I_{2h}^2 A^h w_k^h &= 16c_k s_k w_{2h}^h \quad (36b) \\
\Rightarrow (A^{2h})^{-1} I_{2h}^2 A^h w_k^h &= \frac{16c_k s_k}{4\sin^2 \frac{\kappa s}{\pi}} w_{2h}^k = w_{2h}^k \quad (36c) \\
\Rightarrow -I_{2h}^2 (A^{2h})^{-1} I_{2h}^2 A^h w_k^h &= -c_k w_k^h + s_k w_{k'}^h \quad (36d) \\
\Rightarrow [I - I_{2h}^2 (A^{2h})^{-1} I_{2h}^2 A^h] w_k^h &= s_k w_k^h + s_k w_{k'}^h, \quad (36e)
\end{align}

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

\begin{align}
A^h w_k^h &= 4s_k w_k^h \quad (37a) \\
\Rightarrow I_{2h}^2 A^h w_k^h &= -16c_k s_k w_{2h}^h \quad (37b) \\
\Rightarrow (A^{2h})^{-1} I_{2h}^2 A^h w_k^h &= \frac{-16c_k s_k}{4\sin^2 \frac{\kappa s}{\pi}} w_{2h}^k = -w_{2h}^k \quad (37c) \\
\Rightarrow -I_{2h}^2 (A^{2h})^{-1} I_{2h}^2 A^h w_k^h &= c_k w_k^h - s_k w_{k'}^h \quad (37d) \\
\Rightarrow [I - I_{2h}^2 (A^{2h})^{-1} I_{2h}^2 A^h] w_k^h &= c_k w_k^h + c_k w_{k'}^h. \quad (37e)
\end{align}

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 (36c) and $\lambda_k^{\nu_2}$ for (37e). In contrast, adding post-smoothing incurs a scaling of $\lambda_k^{\nu_2}$ for $w_k^h$ and a scaling of $\lambda_k^{\nu_1}$ for $w_{k'}^h$ in both (36e) and (37e). Hence (35) holds.

Remark 5. (35) can be rewritten as

\begin{equation}
TG \begin{bmatrix}
  w_k \\
  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}
  w_k \\
  w_{k'}
\end{bmatrix} = \begin{bmatrix}
  c_1 & c_2 \\
  c_3 & c_4
\end{bmatrix} \begin{bmatrix}
  w_k \\
  w_{k'}
\end{bmatrix}. \quad (38)
\end{equation}

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

\begin{align}
I_{2h}^h &= c(I_{2h}^2)^T, \quad c \in \mathbb{R}. \quad (39a) \\
I_{2h}^2 A^h I_{2h}^h &= A^{2h}. \quad (39b)
\end{align}

(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 $\mathcal{N}(I_{2h}^h) = \{0\}$.

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 $v^{2h} = \sum_{j=1}^{n} \phi_j^{2h} e_j^{2h}$. 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 $\mathcal{N}(A) = \{x : Ax = 0\}$. The left null space is the null space of $A^T$. \[7\]
Figure 1: The damping coefficients of two-grid correction with weighted Jacobi for $n = 64$. The x-axis is $k$. It is tricky to plot the coefficients defined in (38), especially in Matlab. Since $c_2, c_4$ act on HF modes, one has to ensure that the components in the vectors $s_k$ and $c_k$ indeed correspond to those in $w_k'$. If $s_k$ and $c_k$ are computed from an increasing order of the frequencies, then their components will have to be reversed for plotting. Physical intuition helps in this case: $c_1$ and $c_4$ should form one curve while $c_2$ and $c_3$ should form another.
Theorem 26 (The counting theorem or the fundamental theorem of linear maps). Suppose the vector space $V$ is finite-dimensional and $T \in L(V, W)$. Then the range of $T$ is finite-dimensional and

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

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

\begin{align*}
\mathbb{R}^m &= \mathcal{R}(A) \oplus \mathcal{N}(A^T), \\
\mathbb{R}^n &= \mathcal{R}(A^T) \oplus \mathcal{N}(A),
\end{align*}

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

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

$$\begin{bmatrix}
a_1^T \\
a_2^T \\
\vdots \\
a_m^T
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_m
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0
\end{bmatrix},$$

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

Remark 6. Each $x \in \mathbb{R}^n$ can be split into a row space component $x_r$ and a null space component $x_n$. Then $Ax = Ax_r \in \mathcal{R}(A)$. 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,

$$\dim \mathcal{R}(I^{2h}_h) = \frac{n}{2} - 1, \quad \dim \mathcal{N}(I^{2h}_h) = \frac{n}{2}.$$  \hspace{1cm} (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

$$\begin{align*}
\mathcal{R}(A) &= \text{span}\{U_1, U_2, \ldots, U_r\}, \\
\mathcal{N}(A) &= \text{span}\{V_{n-r+1}, V_{n-r+2}, \ldots, V_n\}, \\
\mathcal{R}(A^T) &= \text{span}\{V_1, V_2, \ldots, V_r\}, \\
\mathcal{N}(A^T) &= \text{span}\{U_{m-r+1}, U_{m-r+2}, \ldots, U_m\}.
\end{align*}$$

This is closely related to Theorem 27.

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

$$\mathcal{N}(I^{2h}_h) = \text{span}\{A^h e_j^h : j \text{ is odd}\},$$

where $e_j^h$ is the $j$th unit vector on $\Omega^h$.

Proof. Consider $I^{2h}_h A^h$. The $j$th row of $I^{2h}_h$ 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 $\sum (I^{2h}_h)_{ij}(A^h)_{ik}$. 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 $I^{2h}_h A^h e_j^h = 0$. □
The above proposition states that the basis vector of $\mathcal{N}(I_h^{2h})$ are of the form

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

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:

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

**Proof.** If $s^h \in \mathcal{R}(I_h^{2h})$, then

$$TGs^h = [I - I_h^{2h}(A^{2h})^{-1}I_h^{2h}A^h]I_h^{2h}q^{2h} = 0,$$

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

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

$$TGt^h = [I - I_h^{2h}(A^{2h})^{-1}I_h^{2h}A^h]t^h = t^h,$$

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 $\dim \mathcal{R}(I_h^{2h})$ since $A^h$ is a bijection with full rank on $\mathbb{R}^{n-1}$. This implies that $\dim \mathcal{N}(TG) \leq \dim \mathcal{R}(I_h^{2h})$, which completes the proof.

Now that we have both the spectral decomposition $\Omega^h = L \oplus H$ and the subspace decomposition $\Omega^h = \mathcal{R}(I_h^{2h}) \oplus \mathcal{N}(I_h^{2h}A^h)$, the combination of relaxation with $TG$ correction is equivalent to projecting the initial error vector to the $L$ axis and then to the $N$ 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

$$v^h \leftarrow V^h(v^h, f^h, v_1, v_2) \quad (48)$$

with the following steps.

1) relax $v_1$ times on $A^h u^h = f^h$ with a given initial guess $v^h$,

2) if $\Omega^h$ is the coarsest grid, go to step 4), otherwise

$$f^{2h} \leftarrow I_h^{2h}(f^h - A v^h),$$

$$v^{2h} \leftarrow 0,$$

$$v^{2h} \leftarrow V^h(v^{2h}, f^{2h}).$$

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

4) relax $v_2$ times on $A^h u^h = f^h$ with the initial guess $v^h$.

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

$$v^h \leftarrow FMG^h(f^h, v_1, v_2) \quad (49)$$

with the following steps.
1) If \( \Omega^h \) is the coarsest grid, set \( v^h \leftarrow 0 \) and go to step 3), otherwise

\[
\begin{align*}
 f^{2h} & \leftarrow I_h^{2h} f^h, \\
v^{2h} & \leftarrow FMG^{2h}(f^{2h}, \nu_1, \nu_2).
\end{align*}
\]

2) correct \( v^h \leftarrow I^h_{2h} v^{2h}, \)

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

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.