Technische Universität München
Lehrstuhl für Informatik V
Univ.-Prof. Dr. M. Bader
Carsten Uphoff, Nikola Tchipev

SS 2018
Sheet 2
27.04.2018

Scientific Computing II

Iterative Solvers

Exercise 1: Fourier Analysis for Two-Grid Iteration

We stick to the Poisson example:

$$
\begin{aligned}
-\frac{d^2u}{dx^2} &= f(x), \quad x \in (0,1), \\
u(0) = u(1) &= 0.
\end{aligned}
$$

This time, however, we want to consider the more general case of having $N+1$ grid points, where $N := 2K$, and a resulting mesh size $h := 1/N$. A step towards multigrid methods consists in the *two-grid method* based on *coarse-grid correction*. In this particular case, we formulate the algorithm as follows:

- After some iteration steps using a smoother (like Jacobi), compute residual $r^{old} := b^h - A_h u^h$. The residual and the error e^{old} fulfill the residual equation $r^{old} = Ae^{old}$.
- Restrict the residual to the coarse grid using an operator R. This yields a (residual-like) vector r^{2h} on the coarse grid. The coarse grid is assumed to have $K+1$ points.
- Solve the residual equation $A_{2h}e^{2h} = r^{2h}$ on the coarse grid. A_{2h} corresponds to the coarse-grained system that resembles A_h on the fine grid.
- Interpolate the resulting coarse-grid approximation of the error e^{2h} to the fine grid and correct the fine-grid error $e^{new} := e^{old} - Pe^{2h}$. The interpolation operator shall be denoted by P.

Furthermore, we define vectors q_h^m with components $(q_h^m)_i := \sin(m\pi hi)$. With respect to the fine grid, we have a low frequency if $m < N/2$, and a high frequency if $m \geq N/2$.
In the following, we want to step through the Fourier analysis for this method and investigate the convergence behaviour, considering one cycle of this two-grid algorithm.

(a) Give a closed expression for e^{new} which only depends on e^{old} and not on e^{2h} anymore. You may use the operators R, P, A_{2h} and A_h from above.

(b) Define the restriction operator R by injection (see lecture slides). Show that the restriction of any error frequency Rq_h^m yields a low frequency, that is

$$
Rq_h^m = \begin{cases} q_{2h}^m & \text{if} \quad m < N/2 \\ -q_{2h}^{N-m} & \text{if} \quad m \geq N/2. \end{cases}
$$

(c) For interpolation, we want to use the linear interpolation scheme:

$$(Pq_{2h}^m)_i = \begin{cases} \underbrace{(q_{2h}^m)_{i/2}}_{= (q_h^m)_i} & \text{for} \quad i = 2, 4, \dots, N-2 \\ \frac{1}{2}\big(\underbrace{(q_{2h}^m)_{(i-1)/2}}_{= (q_h^m)_{i-1}} + \underbrace{(q_{2h}^m)_{(i+1)/2}}_{= (q_h^m)_{i+1}}\big) & \text{for} \quad i = 1, 3, \dots, N-1. \end{cases}$$

Use the function definitions

$$\frac{1}{2}(\cos(\pi i) + 1) = \begin{cases} 1 & \text{for } i = 2, 4, \dots, N-2 \\ 0 & \text{for } i = 1, 3, \dots, N-1 \end{cases}$$

$$\frac{1}{2}(-\cos(\pi i) + 1) = \begin{cases} 0 & \text{for } i = 2, 4, \dots, N-2 \\ 1 & \text{for } i = 1, 3, \dots, N-1 \end{cases}$$

and

$$(q_h^{N-m})_i = \sin((N-m)\pi h i) = \underbrace{\sin(N\pi h i)}_{=0}\cos(m\pi h i) - \cos(N\pi h i)\underbrace{\sin(m\pi h i)}_{=(q_h^m)_i} = -\cos(\pi i)(q_h^m)_i$$

to re-write the interpolated coarse grid frequency Pq_{2h}^m as

$$Pq_{2h}^m = a_m \cdot q_h^m + b_m \cdot q_h^{N-m}$$

with (frequency-dependent) constants a_m, b_m. What does the latter equation tell us about the frequency of the interpolated function Pq_{2h}^m?

(d) Putting the results from prolongation and restriction together, we can derive by further computation (not discussed in the exercise) that

$$\begin{aligned} PA_{2h}^{-1}RA_h q_h^m &= \begin{cases} q_h^m - \frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2} q_h^{N-m} & \text{if} \quad m < \frac{N}{2} \\ -\frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2} q_h^{N-m} + q_h^m & \text{if} \quad m \geq \frac{N}{2} \end{cases} \\ &= q_h^m - \frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2} q_h^{N-m} \end{aligned}$$

Assume that the initial error is a linear combination of two sines: $e^{old} = x_1 q_h^m + x_2 q_h^{N-m}$ with $m < N/2$. Express $PA_{2h}^{-1}RA_h e^{old}$ in terms of q_h^m and q_h^{N-m}, a 2×2-matrix B_m, and the vector $(x_1, x_2)^T$. That is, find a representation

$$PA_{2h}^{-1}RA_h e^{old} = \begin{pmatrix} q_h^m & q_h^{N-m} \end{pmatrix} B_m \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

With this in mind, we can simplify the analysis with a change of basis. Let

$$Q = (q_h^1, q_h^{N-1}, q_h^2, q_h^{N-2}, \dots, q_h^{K-1}, q_h^{N-K+1}, q_h^K).$$

As this matrix is invertible, we can find a vector x^{new} with $e^{new} = Qx^{new}$ and a vector x^{old} with $e^{old} = Qx^{old}$. The entry $(x^{old})_i$ can be interpreted as the weight of the frequency in the i-th column of Q. If we left-multiply e^{new} with Q^{-1} we have

$$x^{new} = Q^{-1}e^{new} = Q^{-1}e^{old} - Q^{-1}PA_{2h}^{-1}RA_he^{old} = x^{old} - Cx^{old} = (I - C)x^{old}.$$

That is, we have found a frequency mapping $x \mapsto (I - C)x$ and may directly work with the error's frequency components instead with the error itself.

Determine the matrix C. (C will be block diagonal and consist of blocks B_m.)

(e) In order to remove high frequency errors, we apply two Jacobi iterations before (*pre-smoothing*) and after (*post-smoothing*) the coarse-grid correction scheme. Following the Fourier analysis (see Sheet 1) for the basis vectors $(q_h^m)_{m=1,...,N-1}$, one weighted Jacobi iteration ($\omega = 1/2$) corresponds to the update rule (again working on the error's frequency components with the same ordering as in (d))

$$x \mapsto M_f x.$$

I.e. for the frequency pair m and $N - m$ we obtain

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{1}{2}(1 + \cos(m\pi h)) & 0 \\ 0 & \frac{1}{2}(1 - \cos(m\pi h)) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Because C is a block diagonal matrix and M_f is diagonal, we may limit the discussion to frequency pairs. The general results then follow easily.

Give an estimate for the maximum eigenvalue of the overall solver operations, that is pre-smoothing $\rightarrow$ restriction $\rightarrow$ coarse-grid correction $\rightarrow$ interpolation $\rightarrow$post-smoothing.

Solution:

(a) The closed expression reads

$$\begin{aligned} e^{new} &= e^{old} - Pe^{2h} = e^{old} - PA_{2h}^{-1} r^{2h} = e^{old} - PA_{2h}^{-1} R r^{old} = e^{old} - PA_{2h}^{-1} R A_h e^{old} \\ &= (I - PA_{2h}^{-1} R A_h) e^{old}. \end{aligned}$$

(b) If we apply restriction by injection, we can only resolve the frequency at every second grid point. This means that we can represent a frequency q_h^m on the coarse grid as

$$Rq_h^m = (\sin(m\pi h i))_{i=2,4,\ldots,N-2} = (\sin(m\pi 2hi))_{i=1,\ldots,\frac{N}{2}-1}. \tag{1}$$

For a low frequency $m < N/2$ which can still be resolved on the coarse grid, we can interprete the latter as frequency q_{2h}^m on the coarse grid nodes.
For higher frequencies $m \geq N/2$, we know that these cannot be resolved on the coarse grid anymore. Still, they should yield a contribution on the coarse grid. To understand their behaviour, we transform Eq. (1) using the sum formula for the sine:

$$\begin{aligned} \sin(m\pi 2hi) &= \sin((m - N + N)\pi 2hi) \\ &= \sin((m-N)\pi 2hi) \underbrace{\cos(N\pi 2hi)}_{=\cos(2\pi i)=1} + \cos((m-N)\pi 2hi) \underbrace{\sin(N\pi 2hi)}_{=\sin(2\pi i)=0} \\ &= -\sin((N-m)\pi 2hi). \end{aligned}$$

The latter term $\sin((N-m)\pi 2hi)$ corresponds to a low frequency again (since for $m \geq N/2$, we have $N - m \leq N/2$). The restriction of a frequency q_h^m thus yields:

$$Rq_h^m = \begin{cases} q_{2h}^m & \text{if} \quad m < \frac{N}{2} \\ -q_{2h}^{N-m} & \text{if} \quad m \geq \frac{N}{2}. \end{cases}$$

(c) The cosine-based function definitions can be used to re-write the case distinction in a single formula using a linear combination of both cases:

$$\begin{aligned} (Pq_{2h}^m)_i &= \tfrac{1}{2}(\cos(\pi i) + 1)(q_h^m)_i + \tfrac{1}{2}(-\cos(\pi i) + 1)\tfrac{1}{2}\left((q_h^m)_{i-1} + (q_h^m)_{i+1}\right) \\ &= \tfrac{1}{2}(\cos(\pi i) + 1)(q_h^m)_i + \tfrac{1}{2}(-\cos(\pi i) + 1)\cos(m\pi h)(q_h^m)_i \\ &= \tfrac{1}{2}\left[(1 + \cos(m\pi h))(q_h^m)_i + (1 - \cos(m\pi h)) \underbrace{\cos(\pi i)}_{=\cos(N\pi hi)} \underbrace{(q_h^m)_i}_{=\sin(m\pi hi)}\right] \end{aligned} \tag{2}$$

The step from the first to the second row of Eq. (2) involves again the sum formula for the sine (see also exercise 2 from the previous sheet). The last row of Eq. (2) only resembles a re-ordering of the terms.
In the last row of Eq. (2), we still have a coefficient $\cos(\pi i)$ which yields a spatial dependency of the frequency after interpolation. We remove this dependency by a similar trick as in the previous discussion of the restriction operator: we therefore consider again the frequency q_h^{N-m} which can be re-written as

$$\underbrace{\sin((N-m)\pi hi)}_{=(q_h^{N-m})_i} = \underbrace{\sin(N\pi hi)}_{=0} \cos(m\pi hi) - \cos(N\pi hi) \underbrace{\sin(m\pi hi)}_{=(q_h^m)_i}.$$

Inserting this relation into Eq. (2) yields

$$(Pq_{2h}^m)_i = \frac{1}{2}\left[(1+\cos(m\pi h))(q_h^m)_i + (-1+\cos(m\pi h))\underbrace{\sin((N-m)\pi hi)}_{(q_h^{N-m})_i}\right]$$

or in vector notation

$$Pq_{2h}^m = \frac{1}{2}\left[(1+\cos(m\pi h))q_h^m + (-1+\cos(m\pi h))q_h^{N-m}\right].$$

From the latter equation, we can observe one interesting thing: if $m < N/2$, i.e. m corresponds to a low frequency, we will obtain a high frequency contribution $N-m$ on the fine grid. If $m \geq N/2$, we will obtain a high frequency contribution m on the fine grid. Consequently, no matter whether the frequency m corresponds to a rather high or low frequency mode, we will always generate a high frequency mode on the fine grid. Hence, after prolongation/ interpolation, we always need to apply *post-smoothing* to remove these generated high frequencies. See also the following table for $N = 8$.

m	a_m	b_m	$N-m$
1	0.9619	-0.0381	7
2	0.8536	-0.1464	6
3	0.6913	-0.3087	5
4	0.5000	-0.5000	4
5	0.3087	-0.6913	3
6	0.1464	-0.8536	2
7	0.0381	-0.9619	1

(d) We can apply $PA_{2h}^{-1}RA_h$ separately to q_h^m and q_h^{N-m} in $e^{old} = x_1 q_h^m + x_2 q_h^{N-m}$ (due to linearity). Hence, we obtain

$$PA_{2h}^{-1}RA_h e^{old} = x_1 q_h^m - x_1\mu_m q_h^{N-m} - x_2\mu_m q_h^m + x_2 q_h^{N-m} = (x_1 - \mu_m x_2)q_h^m + (x_2 - \mu_m x_1)q_h^{N-m}$$

where we introduced $\mu_m = \frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2}$. The latter can be written as a matrix-vector product:

$$\begin{pmatrix} q_h^m & q_h^{N-m} \end{pmatrix}\begin{pmatrix} x_1 - \mu_m x_2 \\ x_2 - \mu_m x_1 \end{pmatrix} = \begin{pmatrix} q_h^m & q_h^{N-m} \end{pmatrix}\underbrace{\begin{pmatrix} 1 & -\mu_m \\ -\mu_m & 1 \end{pmatrix}}_{=B_m}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We see that the operator $PA_{2h}^{-1}RA_h$ does only effect the mode pair q_h^m, q_h^{N-m} and none other. So if we would introduce multiple mode pairs they would not influence each other, such that the matrix C is just a block diagonal matrix:

$$Q\underbrace{\begin{pmatrix} B_1 & & & & \\ & B_2 & & & \\ & & \ddots & & \\ & & & B_{K-1} & \\ & & & & 1-\mu_K \end{pmatrix}}_{=C}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix}$$

(We have to take special care for the case $m = K$, as here $N - m = N - K = K = m$.)

(e) We know how to map frequency components for weighted Jacobi iterations as well as for a two-grid cycle. Hence, we can simply combine those operations:

$$x \mapsto M_f^2 (I - C) M_f^2 x$$

As C is block-diagonal and M_f is diagonal it is sufficient to limit the discussion to frequency pairs and we obtain

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \underbrace{\begin{pmatrix} \frac{1+\cos(m\pi h)}{2} & 0 \\ 0 & \frac{1-\cos(m\pi h)}{2} \end{pmatrix}^2}_{\text{post-smoothing}} \begin{pmatrix} 0 & \frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2} \\ \frac{(1-\cos(m\pi h))^2}{\sin(m\pi h)^2} & 0 \end{pmatrix} \underbrace{\begin{pmatrix} \frac{1+\cos(m\pi h)}{2} & 0 \\ 0 & \frac{1-\cos(m\pi h)}{2} \end{pmatrix}^2}_{\text{pre-smoothing}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

which after some further transformations can be written as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & \frac{\sin(m\pi h)^2(1-\cos(m\pi h))^2}{16} \\ \frac{\sin(m\pi h)^2(1-\cos(m\pi h))^2}{16} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The largest singular value of this matrix is given by

$$\sigma_m = \frac{\sin(m\pi h)^2(1 - \cos(m\pi h))^2}{16}$$

We can thus give an estimate of the largest singular value for all frequency pairs:

$$\sigma_{max} = \max_{1 \le m < N/2} \sigma_m = \max_{1 \le m < N/2} \frac{\sin(m\pi h)^2(1 - \cos(m\pi h))^2}{16} \le \frac{1}{16}.$$

Hence,

$$\|x^{new}\| = \|M_f^2 (I - C) M_f^2 x^{old}\| \le \sigma_{max} \|x^{old}\| \le \frac{1}{16} \|x^{old}\|.$$

Let's review the overall procedure so far: we inserted an error frequency q_h^m into our method, determined the transfer of this frequency from the fine to the coarse grid (restriction) and from the coarse to the fine grid (prolongation/ interpolation). Subsequently, we formulated the overall scheme for this frequency in the subspace of the basis vectors $\{q_h^m, q_h^{N-m}\}$ since this turned out to be sufficient to describe the frequency transfer. After including the pre- and post-smoothing, we could now give an estimate of the eigenvalues of our solver which determine how strong the error is decreased.
From the last equation, we can now see that the error is at least reduced by a factor of 16 in one two-grid cycle. This is even independent from the mesh size (unlike the smoothers discussed so far (such as Jacobi) where we saw that we need the more iterations the more grid points we use). In order to achieve a certain accuracy, this method consequently also does not depend on the resolution h and thus delivers an optimal convergence behaviour.