Advanced Discretization Techniques Homework 8

Bruno Degli Esposti, Xingyu Xu

December 10th - December 17th, 2019

Exercise 17: A FV scheme for a nonlinear equation

a) The vector $\varphi(V)$ is defined as the solution U to the linear system (A), with $h_i f_i(V_i)$ on the right-hand side. Once we replace $\mathcal{F}_{i+1/2}$ and $\mathcal{F}_{i-1/2}$ with their definitions in terms of U_i, we can write the linear system as $AU = b(V)$, where $A \in \mathbb{R}^{N\times N}$ and $b(V) \in \mathbb{R}^N$ are defined as follows:

$$A_{i,i-1} = -\frac{1}{h_{i-1/2}} \ \ \forall\, i = 2,\dots,N, \quad A_{i,i} = \frac{1}{h_{i-1/2}} + \frac{1}{h_{i+1/2}} \ \ \forall\, i = 1,\dots,N,$$

$$A_{i,i+1} = -\frac{1}{h_{i+1/2}} \ \ \forall\, i = 1,\dots,N-1, \quad A_{i,j} = 0 \ \text{ otherwise},$$

$$b(V)_i = h_i f_i(V_i) \quad \forall\, i = 1,\dots,N.$$

The function φ is then defined as $\varphi(V) = A^{-1}b(V)$, provided that A is invertible. By the calculations we will do in point b) (following the hint), we know that

$$U^T A U = \sum_{i=0}^{N} \frac{(U_{i+1} - U_i)^2}{h_{i+1/2}} \geq 0.$$

As long as $U \neq 0$, we will show in point d) that the inequality is strict: this means that A is positive definite. Therefore, A is invertible and φ is well-defined. The right hand side $b(V)$ is not a linear function of V, so $\varphi(V) = A^{-1}b(V)$ cannot be linear. Moving on to continuity, let $\{V^k\}$ be a sequence in $\mathbb{R}^N$ converging to V. We want to prove that $\lim_{k\to+\infty} \varphi(V^k) = \varphi(V)$. By the continuity of A^{-1}, it's enough to show that $\lim_{k\to+\infty} b(V^k)_i = b(V)_i$ holds for each $i = 1,\dots,N$. Indeed, by Lebesgue's dominated convergence lemma and the continuity of $f(x,\cdot)$, we have that

$$\lim_{k\to+\infty} b(V^k)_i = \lim_{k\to+\infty} \int_{\Omega_i} f(x, V_i^k)\, dx = \int_{\Omega_i} \lim_{k\to+\infty} f(x, V_i^k)\, dx$$
$$= \int_{\Omega_i} f(x, V_i)\, dx = b(V)_i.$$

In order to apply Lebesgue's lemma, we can choose $g(x) \equiv \|f\|_\infty$ as the dominating function, clearly integrable in the limited interval Ω_i. Lastly, the fixed-point property is proven as follows:

$$\varphi(U) = U \iff A^{-1}b(U) = U \iff AU = b(U) \iff \text{(A)},$$

by the definitions of A and b.

b) Following the hint, we get that

$$\begin{aligned}
\sum_{i=1}^{N} U_i h_i f_i(V_i) &= \sum_{i=1}^{N} U_i \left(-\frac{U_{i+1} - U_i}{h_{i+1/2}} + \frac{U_i - U_{i-1}}{h_{i-1/2}} \right) \\
&= \sum_{i=1}^{N} \frac{-U_{i+1}U_i + U_i^2}{h_{i+1/2}} + \sum_{i=1}^{N} \frac{U_i^2 - U_{i-1}U_i}{h_{i-1/2}} \\
&= \sum_{i=1}^{N} \frac{-U_{i+1}U_i + U_i^2}{h_{i+1/2}} + \sum_{j=0}^{N-1} \frac{U_{j+1}^2 - U_j U_{j+1}}{h_{j+1/2}} \\
&= \sum_{i=0}^{N} \frac{-U_{i+1}U_i + U_i^2}{h_{i+1/2}} + \sum_{j=0}^{N} \frac{U_{j+1}^2 - U_j U_{j+1}}{h_{j+1/2}} \\
&= \sum_{i=0}^{N} \frac{U_{i+1}^2 - 2U_{i+1}U_i + U_i^2}{h_{i+1/2}} = \sum_{i=0}^{N} \frac{(U_{i+1} - U_i)^2}{h_{i+1/2}}.
\end{aligned}$$

The boundary conditions $U_0 = U_{N+1} = 0$ allowed us to include $i = 0$ and $j = N$ in the range of the sums. Now we can easily complete the proof:

$$\begin{aligned}
\sum_{i=0}^{N} \frac{(U_{i+1} - U_i)^2}{h_{i+1/2}} &= \sum_{i=1}^{N} U_i h_i f_i(V_i) \le \sum_{i=1}^{N} h_i \left| U_i \right| \left| f_i(V_i) \right| \\
&= \sum_{i=1}^{N} h_i \left| U_i \right| \left| \frac{1}{h_i} \int_{\Omega_i} f(x, V_i)\, dx \right| \le \sum_{i=1}^{N} \left| U_i \right| \left| \Omega_i \right| \left\| f \right\|_\infty \\
&= M \sum_{i=1}^{N} (x_{i+1/2} - x_{i-1/2}) \left| U_i \right| = M \sum_{i=1}^{N} h_i \left| U_i \right|.
\end{aligned}$$

c) The boundary condition $U_0 = 0$ allows us to write U_i as a telescopic series. Then, by Cauchy-Schwarz's inequality, we have that

$$\begin{aligned}
|U_i| &= \left| \sum_{j=0}^{i-1} (U_{j+1} - U_j) \right| = \left| \sum_{j=0}^{i-1} \frac{U_{j+1} - U_j}{(h_{j+1/2})^{1/2}} (h_{j+1/2})^{1/2} \right| \\
&\le \left| \sum_{j=0}^{i-1} \frac{(U_{j+1} - U_j)^2}{h_{j+1/2}} \right|^{1/2} \left| \sum_{j=0}^{i-1} h_{j+1/2} \right|^{1/2} \\
&\le \left(\sum_{j=0}^{N} \frac{(U_{j+1} - U_j)^2}{h_{j+1/2}} \right)^{1/2} \left(\sum_{j=0}^{N} h_{j+1/2} \right)^{1/2} = \left(\sum_{j=0}^{N} \frac{(U_{j+1} - U_j)^2}{h_{j+1/2}} \right)^{1/2}.
\end{aligned}$$

Now we can substitute this inequality into the one we've proved in point b):

$$\sum_{i=0}^{N} \frac{(U_{i+1}-U_i)^2}{h_{i+1/2}} \le M \sum_{i=1}^{N} h_i \left|U_i\right| \le M \sum_{i=1}^{N} h_i \left(\sum_{j=0}^{N} \frac{(U_{j+1}-U_j)^2}{h_{j+1/2}} \right)^{1/2}$$

$$\left(\sum_{i=0}^{N} \frac{(U_{i+1}-U_i)^2}{h_{i+1/2}} \right)^{1/2} \le M \sum_{i=1}^{N} h_i = M.$$

Thus we can choose $C(f) = M^2 = \|f\|_\infty^2$.

d) First we prove that

$$\|V\| = \left(\sum_{j=0}^{N} \frac{(V_{j+1}-V_j)^2}{h_{j+1/2}} \right)^{1/2}$$

defines a norm. For each $V \in \mathbb{R}^N$, $\|V\| \ge 0$ because $\|V\|$ is the square root of the sum of positive terms. As for the triangle inequality,

$$\begin{aligned}
\|V+W\|^2 &= \sum_{j=0}^{N} \frac{(V_{j+1}+W_{j+1}-V_j-W_j)^2}{h_{j+1/2}} = \sum_{j=0}^{N} \frac{(V_{j+1}-V_j+W_{j+1}-W_j)^2}{h_{j+1/2}} \\
&= \sum_{j=0}^{N} \frac{(V_{j+1}-V_j)^2+(W_{j+1}-W_j)^2+2(V_{j+1}-V_j)(W_{j+1}-W_j)}{h_{j+1/2}} \\
&\le \|V\|^2 + \|W\|^2 + \sum_{j=0}^{N} \frac{|2(V_{j+1}-V_j)(W_{j+1}-W_j)|}{\sqrt{h_{j+1/2}}\sqrt{h_{j+1/2}}} \\
&\le \|V\|^2 + \|W\|^2 + 2\left(\sum_{j=0}^{N} \frac{(V_{j+1}-V_j)^2}{h_{j+1/2}} \right)^{1/2} \left(\sum_{j=0}^{N} \frac{(W_{j+1}-W_j)^2}{h_{j+1/2}} \right)^{1/2} \\
&= \|V\|^2 + \|W\|^2 + 2\,\|V\|\,\|W\| = (\|V\|+\|W\|)^2.
\end{aligned}$$

Now we check that the norm is absolutely homogeneous:

$$\begin{aligned}
\|\alpha V\| &= \left(\sum_{j=0}^{N} \frac{(\alpha V_{j+1}-\alpha V_j)^2}{h_{j+1/2}} \right)^{1/2} = \left(\sum_{j=0}^{N} \frac{\alpha^2(V_{j+1}-V_j)^2}{h_{j+1/2}} \right)^{1/2} \\
&= \left(\alpha^2 \sum_{j=0}^{N} \frac{(V_{j+1}-V_j)^2}{h_{j+1/2}} \right)^{1/2} = |\alpha| \left(\sum_{j=0}^{N} \frac{(V_{j+1}-V_j)^2}{h_{j+1/2}} \right)^{1/2}.
\end{aligned}$$

Lastly, we need to prove that $\|V\| = 0$ implies $V = 0$. Indeed,

$$
\begin{aligned}
\left(\sum_{j=0}^{N} \frac{(V_{j+1} - V_j)^2}{h_{j+1/2}} \right)^{1/2} = 0 \quad &\Longrightarrow \quad V_{j+1} - V_j = 0 \quad \forall\, j = 0, \ldots, N \\
&\Longrightarrow \quad V_0 = V_1 = \cdots = V_N = V_{N+1} = 0 \\
&\Longrightarrow \quad V = 0.
\end{aligned}
$$

Moving on to the second part of the proof, let $D = \{V \in \mathbb{R}^N \mid \|V\| \leq M\}$ and let $\hat{\varphi} = \varphi|_D$. By point c), we know that the range of $\hat{\varphi}$ is a subset of D. Therefore, the continuous function $\hat{\varphi}\colon D \to D$ has a fixed point U by Brouwer's theorem (which we can apply here, because D is a compact and convex subset of $\mathbb{R}^N$). Then $U = \hat{\varphi}(U) = \varphi(U)$, so the linear system (A) has at least U as a solution by what we've proved in point a). $\square$