∼

Discretization Matrix properties

Scientific Computing Winter 2016/2017

Lecture 23

Jürgen Fuhrmann

juergen.fuhrmann@wias-berlin.de

made wit pandoc

∼

Recap

Time dependent Robin boundary value problem

- Choose final time $T > 0$. Regard functions $(x, t) \to \mathbb{R}$.

$$
\begin{aligned}
\partial_t u - \nabla \cdot \kappa \nabla u = f \quad & \text{in } \Omega \times [0, T] \\
\kappa \nabla u \cdot \vec{n} + \alpha(u - g) = 0 \quad & \text{on } \partial\Omega \times [0, T] \\
u(x, 0) = u_0(x) \quad & \text{in} \Omega
\end{aligned}
$$

- This is an initial boundary value problem
- This problem has a weak formulation in the Sobolev space $L^2\left([0, T], H^1(\Omega)\right)$, which then allows for a Galerkin approximation in a corresponding subspace
- We will proceed in a simpler manner: first, perform a finite difference discretization in time, then perform a finite element (finite volume) discretization in space.
  - Rothe method: first discretize in time, then in space
  - Method of lines: first discretize in space, get a huge ODE system

Time discretization

- Choose time discretization points $0 = t_0 < t_1 \dots < t_N = T$, let
  $\tau_i = t_i - t_{i-1}$
  For $i = 1 \dots N$, solve

$$\frac{u_i - u_{i-1}}{\tau_i} - \nabla \cdot \kappa \nabla u_\theta = f \quad \text{in } \Omega \times [0, T]$$
$$\kappa \nabla u_\theta \cdot \vec{n} + \alpha(u_\theta - g) = 0 \quad \text{on } \partial\Omega \times [0, T]$$

  where $u_\theta = \theta u_i + (1 - \theta) u_{i-1}$
- $\theta = 1$: backward (implicit) Euler method
- $\theta = \frac{1}{2}$: Crank-Nicolson scheme
- $\theta = 0$: forward (explicit) Euler method
- Note that the explicit Euler method does not involve the solution of a PDE system. What do we have to pay for this ?

Weak formulation

- Weak formulation: search $u \in H^1(\Omega)$ such that

$$
\begin{aligned}
\frac{1}{\tau_i}\int_\Omega u_i v\, dx + \theta\left(\int_\Omega \kappa\nabla u_i \nabla v\, dx + \int_{\partial\Omega} \alpha u_i v\; ds\right) = \\
\frac{1}{\tau_i}\int_\Omega u_{i-1} v\, dx + (1-\theta)\left(\int_\Omega \kappa\nabla u_{i-1} \nabla v\, dx + \int_{\partial\Omega} \alpha u_{i-1} v\; ds\right) \\
+ \int_\Omega fv\, dx + \int_{\partial\Omega} \alpha g v\, ds\; \forall v \in H^1(\Omega)
\end{aligned}
$$

- Matrix formulation (in case of constant coefficents, $A_i = A$)

$$
\frac{1}{\tau_i} M u_i + \theta A_i u_i = \frac{1}{\tau_i} M u_{i-1} + (1-\theta) A_i u_{i-1} + F
$$

- M: mass matrix, A: stiffness matrix

$\sim$

Matrix properties

Mass matrix

- Mass matrix $M = (m_{ij})$:
$$m_{ij} = \int_\Omega \phi_i \phi_j \; dx$$
- Self-adjoint, coercive bilinear form $\Rightarrow$ M is symmetric, positiv definite
- For a family of quasi-uniform, shape-regular triangulations, for every eigenvalue μ one has the estimate
$$c_1 h^d \leq \mu \leq c_2 h^d$$
Therefore the condition number $\kappa(M)$ is bounded by a constant independent of h:
$$\kappa(M) \leq c$$
- How to see this ? Let $u_h = \sum_{i=1}^N U_i \phi_i$, and μ an eigenvalue (positive,real!) Then
$$||u_h||_0^2 = (U, MU)_{\mathbb{R}^N} = \mu(U, U)_{\mathbb{R}^N} = \mu ||U||_{\mathbb{R}^N}^2$$
From quasi-uniformity we obtain
$$c_1 h^d ||U||_{\mathbb{R}^N}^2 \leq ||u_h||_0^2 \leq c_2 h^d ||U||_{\mathbb{R}^N}^2$$
and conclude

Mass matrix M-Property ?

- For P^1-finite elements, all integrals $m_{ij} = \int_\Omega \phi_i \phi_j \, dx$ are zero or positive, so we get positive off diagonal elements.
- No M-Property!

Stiffness matrix condition number + row sums

- Stiffness matrix $A = (a_{ij})$:

$$a_{ij} = a(\phi_i, \phi_j) = \int_\Omega \nabla\phi_i \nabla\phi_j \; dx$$

- bilinear form $a(\cdot,\cdot)$ is self-adjoint, therefore A is symmetric, positive definite
- Condition number estimate for P^1 finite elements on quasi-uniform triangulation:

$$\kappa(A) \leq ch^{-2}$$

- Row sums:

$$\begin{aligned}\sum_{j=1}^N a_{ij} &= \sum_{j=1}^N \int_\Omega \nabla\phi_i \nabla\phi_j \; dx = \int_\Omega \nabla\phi_i \nabla\left(\sum_{j=1}^N \phi_j\right) \; dx \\ &= \int_\Omega \nabla\phi_i \nabla\left(1\right) \; dx \\ &= 0\end{aligned}$$

Stiffness matrix entry signs

Local stiffness matrices

$$s_{ij} = \int_K \nabla\lambda_i \nabla\lambda_j \; dx = \frac{|K|}{2|K|^2} \left(y_{i+1} - y_{i+2}, x_{i+2} - x_{i+1}\right) \begin{pmatrix} y_{j+1} - y_{j+2} \\ x_{j+2} - x_{j+1} \end{pmatrix}$$

- Main diagonal entries must be positive
- Local contributions from element stiffness matrices: Scalar products of vectors orthogonal to edges. These are nonpositive if the angle between the edges are $\leq 90^\circ$
- *weakly acute triangulation*: all triangle angles are less than $\leq 90^\circ$
- in fact, for constant coefficients, in *2D*, Delaunay is sufficient!
- All rows sums are zero $\Rightarrow A$ is singular
- Matrix becomes irreducibly diagonally dominant if we add at least one positive value to the main diagonal, e.g. from Dirichlet BC or *lumped* mass matrix $\Rightarrow M - Matrix$
- Adding a mass matrix yields a positive definite matrix and thus nonsingularity, but *destroys* M-property

Back to time dependent problem

Assume M diagonal, $A = S + D$, where S is the stiffness matrix, and D is a nonnegative diagonal matrix. We have

$$
\begin{aligned}
(Su)_i = \sum_j s_{ij} u_j &= s_{ii} u_i + \sum_{i \neq j} s_{ij} u_j \\
&= (-\sum_{i \neq j} s_{ij}) u_i + \sum_{i \neq j} s_{ij} u_j \\
&= \sum_{i \neq j} -s_{ij} (u_i - u_j)
\end{aligned}
$$

Forward Euler

$$\frac{1}{\tau_i} M u_i = \frac{1}{\tau_i} M u_{i-1} + A_i u_{i-1}$$
$$u_i = u_{i-1} + \tau_i M^{-1} A_i u_{i-1} = (I + \tau M^{-1} D + \tau M^{-1} S) u_{i-1}$$

- Entries of $\tau M^{-1} A) u_{i-1}$ are of order $\frac{1}{h^2}$, and so we can expect stabilityonly if τ balances $\frac{1}{h^2}$, i.e.
$$\tau \leq Ch^2$$
- A more thorough stability estimate proves this situation

Backward Euler

$$\begin{aligned}
\frac{1}{\tau_i} M u_i + A u_i &= \frac{1}{\tau_i} M u_{i-1} \\
(I + \tau_i M^{-1} A) u_i &= u_{i-1} \\
u_i &= (I + \tau_i M^{-1} A)^{-1} u_{i-1}
\end{aligned}$$

But here, we can estimate that

$$||(I + \tau_i M^{-1} A)^{-1}||_\infty \leq 1$$

Backward Euler Estimate

Theorem: Assume S has the sign pattern of an M-Matrix with row sum zero, and D is a nonnegative diagonal matrix. Then $||(I+D+S)^{-1}||_\infty \leq 1$

Proof: Assume that $||(I+S)^{-1}||_\infty > 1$. We know that $(I+S)^{-1}$ has positive entries. Then for α_{ij} being the entries of $(I+S)^{-1}$,

$$max_{i=1}^n \sum_{j=1}^n \alpha_{ij} > 1.$$

Let k be a row where the maximum is reached. Let $e = (1\dots 1)^T$. Then for $v = (I+S)^{-1}e$ we have that $v > 0$, $v_k > 1$ and $v_k \geq v_j$ for all $j \neq k$. The kth equation of $e = (I+S)v$ then looks like

$$\begin{aligned} 1 &= v_k + v_k \sum_{j\neq k} |s_{kj}| - \sum_{j\neq k} |s_{kj}| v_j \\ &\geq v_k + v_k \sum_{j\neq k} |s_{kj}| - \sum_{j\neq k} |s_{kj}| v_k \\ &= v_k > 1 \end{aligned}$$

This contradiction enforces $||(I+S)^{-1}||_\infty \leq 1$.

Backward Euler Estimate II

$$
\begin{aligned}
I + A =& I + D + S \\
=& (I + D)(I + D)^{-1}(I + D + S) \\
=& (I + D)(I + A_{D0})
\end{aligned}
$$

with $A_{D0} = (I + D)^{-1}S$ has row sum zero Thus

$$
\begin{aligned}
||(I + A)^{-1}||_\infty =& ||(I + A_{D0})^{-1}(I + D)^{-1}||_\infty \\
\leq& ||(I + D)^{-1}||_\infty \\
\leq& 1,
\end{aligned}
$$

because all main diagonal entries of $I + D$ are greater or equal to 1. $\square$

Backward Euler Estimate III

We can estimate that

$$I + \tau_i M^{-1} A = I + \tau_i M^{-1} D + \tau_i M^{-1} S$$

and obtain

$$||(I + \tau_i M^{-1} A)^{-1}||_\infty \leq 1$$

- We get this stability independent of the time step.
- Another theory is possible using L^2 estimates and positive definiteness

Discrete maximum principle

Assuming $v \geq 0$ we can conclude $u \geq 0$.

$$\frac{1}{\tau}Mu + (D+S)u = \frac{1}{\tau}Mv$$

$$(\tau m_i + d_i)u_i + s_{ii}u_i = \tau m_i v_i + \sum_{i \neq j}(-s_{ij})u_j$$

$$u_i = \frac{1}{\tau m_i + d_i + \sum_{i \neq j}(-s_{ij})}(\tau m_i v_i + \sum_{i \neq j}(-s_{ij})u_j)$$

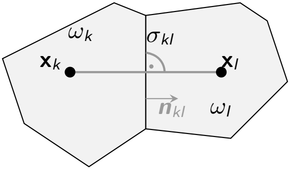
$$\leq \frac{\tau m_i v_i + \sum_{i \neq j}(-s_{ij})u_j}{\tau m_i + d_i + \sum_{i \neq j}(-s_{ij})} \max(\{v_i\} \cup \{u_j\}_{j \neq i})$$

$$\leq \max(\{v_i\} \cup \{u_j\}_{j \neq i})$$

- Provided, the right hand side is zero, the solution in a given node is bounded by the value from the old timestep, and by the solution in the neigboring points.
- No new local maxima can appear during time evolution
- There is a continuous counterpart which can be derived from weak solution
- M-property is crucial for the proof.

The finite volume idea revisited

- Assume Ω is a polygon
- Subdivide the domain Ω into a finite number of **control volumes** :
  $\bar{\Omega} = \bigcup_{k \in \mathcal{N}} \bar{\omega}_k$
  such that
  - ω_k are open (not containing their boundary) convex domains
  - $\omega_k \cap \omega_l = \emptyset$ if $\omega_k \neq \omega_l$
  - $\sigma_{kl} = \bar{\omega}_k \cap \bar{\omega}_l$ are either empty, points or straight lines
    - we will write $|\sigma_{kl}|$ for the length
    - if $|\sigma_{kl}| > 0$ we say that ω_k, ω_l are neigbours
    - neigbours of ω_k: $\mathcal{N}_k = \{l \in \mathcal{N} : |\sigma_{kl}| > 0\}$
- To each control volume ω_k assign a **collocation point**: $\mathbf{x}_k \in \bar{\omega}_k$ such that
  - **admissibility condition**: if $l \in \mathcal{N}_k$ then the line $\mathbf{x}_k\mathbf{x}_l$ is orthogonal to σ_{kl}
  - if ω_k is situated at the boundary, i.e. $\gamma_k = \partial\omega_k \cap \partial\Omega \neq \emptyset$, then $\mathbf{x}_k \in \partial\Omega$



- Now, we know how to construct this partition
  - obtain a boundary conforming Delaunay triangulation
  - construct restricted Voronoi cells

Finite volumes for time dependent problem

Search function $u : \Omega \times [0, T] \to \mathbb{R}$ such that $u(x, 0) = u_0(x)$ and

$$
\begin{aligned}
\partial_t u - \nabla \cdot \lambda \nabla u &= 0 \quad \text{in} \Omega \times [0, T] \\
\lambda \nabla u \cdot \mathbf{n} + \alpha(u - w) &= 0 \quad \text{on} \Gamma \times [0, T]
\end{aligned}
$$

- Given control volume ω_k, integrate equation over space-time control volume

$$
\begin{aligned}
0 &= \int_{\omega_k} \left(\frac{1}{\tau}(u - v) - \nabla \cdot \lambda \nabla u \right) d\omega = -\int_{\partial\omega_k} \lambda \nabla u \cdot \mathbf{n}_k d\gamma + \frac{1}{\tau} \int_{\omega_k} (u - v) d\omega \\
&= -\sum_{L \in \mathcal{N}_k} \int_{\sigma_{kl}} \lambda \nabla u \cdot \mathbf{n}_{kl} d\gamma - \int_{\gamma_k} \lambda \nabla u \cdot \mathbf{n} d\gamma - \frac{1}{\tau} \int_{\omega_k} (u - v) d\omega \\
&\approx \frac{|\omega_k|}{\tau}(u_k - v_k) + \sum_{L \in \mathcal{N}_k} \frac{|\sigma_{kl}|}{h_{kl}}(u_k - u_l) + |\gamma_k| \alpha (u_k - w_k)
\end{aligned}
$$

- Here, $u_k = u(\mathbf{x}_k)$, $w_k = w(\mathbf{x}_k)$, $f_k = f(\mathbf{x}_k)$
- $\frac{1}{\tau_i} M u_i + A u_i = \frac{1}{\tau_i} M u_{i-1}$

$\sim$

Convection-Diffusion

The convection - diffusion equation

Search function $u : \Omega \times [0, T] \to \mathbb{R}$ such that $u(x, 0) = u_0(x)$ and

$$\partial_t u - \nabla(\cdot D\nabla u - u\mathbf{v}) = 0 \quad \text{in}\Omega \times [0, T]$$
$$(D\nabla u - u\mathbf{v}) \cdot \mathbf{n} + \alpha(u - w) = 0 \quad \text{on}\Gamma \times [0, T]$$

- Here:
  - u: species concentration
  - D: diffusion coefficient
  - $\mathbf{v}$: velocity of medium (e.g. fluid)

$$\frac{|\omega_k|}{\tau}(u_k - v_k) + \sum_{L \in \mathcal{N}_k} \frac{|\sigma_{kl}|}{h_{kl}} g(u_k, u_l) + |\gamma_k|\alpha(u_k - w_k)$$

Let $v_{kl} = \frac{1}{|\sigma_{kl}|} \int \sigma_{kl}\mathbf{v} \cdot \mathbf{n}_{kl} d\gamma$

Finite volumes for convection - diffusion II

- Central difference flux:

$$
\begin{aligned}
g(u_k, u_l) &= D(u_k - u_l) - h_{kl}\frac{1}{2}(u_k + u_l)v_{kl} \\
&= (D - \frac{1}{2}h_{kl}v_{kl})u_k - (D + \frac{1}{2}h_{kl}v_{kl})xu_l
\end{aligned}
$$

  - M-Property (sign pattern) only guaranteed for $h \to 0$!
- Upwind flux:

$$
\begin{aligned}
g(u_k, u_l) &= D(u_k - u_l) + \begin{cases} h_{kl}u_k v_{kl}, & v_{kl} < 0 \\ h_{kl}u_l v_{kl}, & v_{kl} > 0 \end{cases} \\
&= (D + \tilde{D})(u_k - u_l) - h_{kl}\frac{1}{2}(u_k + u_l)v_{kl}
\end{aligned}
$$

  - M-Property guaranteed unconditonally !
  - Artificial diffusion $\tilde{D} = \frac{1}{2}h_{kl}|v_{kl}|$

Finite volumes for convection - diffusion: exponential fitting

Project equation onto edge $x_K x_L$ of length $h = h_{kl}$, integrate once - $q = -v_{kl}$

$$
\begin{aligned}
c' + cq &= j \\
c|_0 &= c_K \\
c|_h &= c_L
\end{aligned}
$$

Solution of the homogeneus problem:

$$
\begin{aligned}
c' &= -cq \\
c'/c &= -q \\
\ln c &= c_0 - qx \\
c &= K \exp(-qx)
\end{aligned}
$$

Exponential fitting II

Solution of the inhomogeneous problem: set $K = K(x)$:

$$
\begin{aligned}
K' \exp(-qx) - qK \exp(-qx) + qK \exp(-qx) &= j \\
K' &= j \exp(qx) \\
K &= K_0 + \frac{1}{q} j \exp(qx)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
c &= K_0 \exp(-qx) + \frac{1}{q} j \\
c_K &= K_0 + \frac{1}{q} j \\
c_L &= K_0 \exp(-qh) + \frac{1}{q} j
\end{aligned}
$$

Exponential fitting III

Use boundary conditions

$$
\begin{aligned}
K_0 &= \frac{c_K - c_L}{1-\exp(-qh)} \\
c_K &= \frac{c_K - c_L}{1-\exp(-qh)} + \frac{1}{q} j \\
j =& q c_K - \frac{q}{1-\exp(-qh)}(c_K - c_L) \\
=& q(1 - \frac{1}{1-\exp(-qh)}) c_K - \frac{q}{\exp(-qh)-1} c_L \\
=& q(\frac{-\exp(-qh)}{1-\exp(-qh)}) c_K - \frac{q}{\exp(-qh)-1} c_L \\
=& \frac{-q}{\exp(qh)-1} c_K - \frac{q}{\exp(-qh)-1} c_L \\
=& \frac{B(-qh)c_L - B(qh)c_K}{h}
\end{aligned}
$$

where $B(\xi) = \frac{\xi}{\exp(\xi)-1}$: Bernoulli function

Exponential fitting IV

- Upwind flux:

$$g(u_k, u_l) = D(B(\frac{-v_{kl}h_{kl}}{D})u_k - B(\frac{v_{kl}h_{kl}}{D})u_l)$$

- Allen+Southwell 1955
- Scharfetter+Gummel 1969
- Ilin 1969
- Chang+Cooper 1970
- Guaranteed M property!