Lectures Références

Roger Peyret (NICE ESSI : 89),
Tim Warburton (Boston MIT : 03-05),
Pierre Charrier (Bordeaux Matmeca 96-08)

Numerical Methods for PDE: Finite Differences and Finites Volumes

B. Nkonga

JAD/INRIA

2009

Plan

Plan

Plan

Plan

Plan

1D Poisson Equation

$$-\left(\frac{\partial^2 T}{\partial \mathbf{x}^2}\right) = S(\mathbf{x}), \quad \forall \mathbf{x} \in (0,1), \text{ with } T(0) = T(1) = 0$$

$$T(x) = \int_0^1 \mathcal{G}(x,y) S(y) dy$$

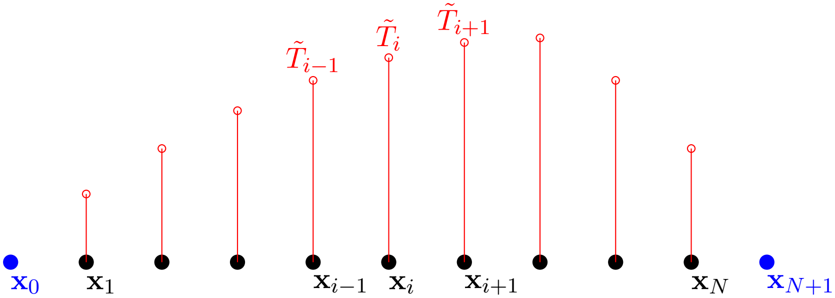
where the Green's function is defined as :

$$\mathcal{G}(x,y) = \begin{cases} y(1-x) & \text{if } 0 \leq y \leq x \\ x(1-y) & \text{if } x \leq y \leq 1 \end{cases}$$

1D Poisson equation : properties

- Existence : The solution $T(\mathbf{x})$ always exists and is unique.
- Regularity : $T(\mathbf{x})$ is always smoother than $S(\mathbf{x})$.
- Positivity : If $S(\mathbf{x}) \geq 0$ for all $\mathbf{x}$ then $T(\mathbf{x}) \geq 0$ for all $\mathbf{x}$.
- Maximum principle : $\|T\|_\infty \leq \frac{1}{8}\|S\|_\infty$

1D mesh for the discretization (approximated solution)



$$\mathbf{x}_0 = 0, \quad \mathbf{x}_{N+1} = 1, \quad \mathbf{x}_i = i\delta x, \quad \delta x = \frac{1}{N+1}, \quad \tilde{T}_i = \tilde{T}(\mathbf{x}_i)$$

Numerical Scheme : Reduced the initial BVP to the computation of the unknowns $\tilde{T}_i$ for $i = 1, ..., N$.

Question : How !

Finite difference strategy : General concept

- Use Taylor's expansions :

$$A(\mathbf{x}+\beta\delta\mathbf{x}) = A(\mathbf{x})+\beta\delta\mathbf{x}\frac{dA}{d\mathbf{x}}(\mathbf{x})+\beta^2\frac{\delta\mathbf{x}^2}{2}\frac{d^2A}{d\mathbf{x}^2}(\mathbf{x})+...+\beta^m\frac{\delta\mathbf{x}^m}{m!}\frac{d^mA}{d\mathbf{x}^m}(\mathbf{x})+R_m(\mathbf{x}$$

- for appropriate set $\vartheta \subset \mathbb{Z}$ of β
- combine them to defined the needed derivative $\frac{d^sA}{d\mathbf{x}^s}(\mathbf{x})$ by eliminating the previous derivatives $1 \leq m < s$.
- do not consider derivatives of order $> s$ and use this truncated formula for mesh points $\mathbf{x} = \mathbf{x}_i$ with unknowns.

Finite difference strategy : Application to $-\frac{d^2T}{d\mathbf{x}^2}(\mathbf{x}) = S(\mathbf{x})$

- Let us choose the set $\vartheta = \{-1, 1\}$.
- We makes Taylor's expansion for values of this set :

$$T(\mathbf{x}+\delta\mathbf{x}) = T(\mathbf{x}) + \delta\mathbf{x}\frac{dT}{d\mathbf{x}}(\mathbf{x}) + \frac{\delta\mathbf{x}^2}{2}\frac{d^2T}{d\mathbf{x}^2}(\mathbf{x}) + R_2(\mathbf{x},\delta\mathbf{x})$$

$$T(\mathbf{x}-\delta\mathbf{x}) = T(\mathbf{x}) - \delta\mathbf{x}\frac{dT}{d\mathbf{x}}(\mathbf{x}) + \frac{\delta\mathbf{x}^2}{2}\frac{d^2T}{d\mathbf{x}^2}(\mathbf{x}) + R_2(\mathbf{x},-\delta\mathbf{x})$$

- Elimination of the first order derivative :
  $T(\mathbf{x}+\delta\mathbf{x}) + T(\mathbf{x}-\delta\mathbf{x}) =$
  $2\tilde{T}(\mathbf{x}) + \delta\mathbf{x}^2\frac{d^2\tilde{T}}{d\mathbf{x}^2}(\mathbf{x}) + R_2(\mathbf{x},\delta\mathbf{x}) + R_2(\mathbf{x},-\delta\mathbf{x})$
- Truncated formula $\delta\mathbf{x}^2\mathcal{E}(\mathbf{x}) = R_2(\mathbf{x},\delta\mathbf{x}) + R_2(\mathbf{x},-\delta\mathbf{x})$:
  $\tilde{T}(\mathbf{x}+\delta\mathbf{x}) + \tilde{T}(\mathbf{x}-\delta\mathbf{x}) = 2\tilde{T}(\mathbf{x}) + \delta\mathbf{x}^2\frac{d^2\tilde{T}}{d\mathbf{x}^2}(\mathbf{x})$
- At the mesh point x_i :
  $\tilde{T}_{i+1} - 2\tilde{T}_i + \tilde{T}_{i-1} = \delta\mathbf{x}^2\frac{d^2\tilde{T}}{d\mathbf{x}^2}(\mathbf{x}) \simeq -\delta\mathbf{x}^2 S_i$

  FD scheme : $-\dfrac{\tilde{T}_{i+1} - 2\tilde{T}_i + \tilde{T}_{i-1}}{\delta\mathbf{x}^2} = S_i$ for $i = 1, ..., N$

Finite difference strategy : Numerical scheme

$$\underline{\mathcal{A}}\tilde{\boldsymbol{T}} = \boldsymbol{S}$$

where

$$\underline{\mathcal{A}} = \frac{1}{\delta \mathbf{x}^2}\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}, \quad \tilde{\boldsymbol{T}} = \begin{pmatrix} \tilde{T}_1 \\ \vdots \\ \tilde{T}_i \\ \vdots \\ \tilde{T}_N \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} S_1 \\ \vdots \\ S_i \\ \vdots \\ S_N \end{pmatrix}$$

Numerical scheme : properties and consequences

- $\underline{\mathcal{A}}$ is symmetric.
- $\underline{\mathcal{A}}$ is non singular : $(\boldsymbol{T}, \underline{\mathcal{A}}\boldsymbol{T}) > 0$ for any $\boldsymbol{T} \neq 0 \in \mathbb{R}^N$. Indeed

$$(\boldsymbol{T}, \underline{\mathcal{A}}\boldsymbol{T}) = \frac{1}{\delta \mathbf{x}^2}\left(T_1^2 + \sum_{i=2}^{N}(T_i - T_{i-1})^2 + \tilde{T}_N^2\right)$$

  Therefore the numerical solution $\tilde{\boldsymbol{T}}$ exists and is unique.
- $\underline{\mathcal{A}}$ is diagonal dominant : $|\mathcal{A}_{ii}| \geq \sum_{j\neq i} |\mathcal{A}_{ij}|$ for all i,
- $\underline{\mathcal{A}}^{-1}$ is not a M-matrix.
  Indeed a M-matrix satisfies
  - for all $i = 1, ...N$: $\mathcal{A}_{ii} > 0$ and $\sum_{j=1}^{N} \mathcal{A}_{ii} > 0$
  - for all $j \neq i$: $\mathcal{A}_{ij} \leq 0$.

Numerical scheme : properties and consequences

$\underline{A}^{-1}$ is non negative : this means that $\left(\underline{A}^{-1}\right)_{ij} \geq 0$.

It is also equivalent to a maximum principle :

$$\text{for any} \quad \boldsymbol{S} \geq 0, \quad \text{if} \quad \underline{A}\tilde{\boldsymbol{T}} = \boldsymbol{S} \quad \text{therefore} \quad \tilde{\boldsymbol{T}} \geq 0$$

where $\boldsymbol{S} \geq 0$ means that $S_i \geq 0$ for $i = 1, ...N$.

Indeed : if $\boldsymbol{S} \geq 0$ then, for $i = 2, ...N-1$ we have

$$-\tilde{T}_{i+1} + 2\tilde{T}_i - \tilde{T}_{i-1} \geq 0 \Longrightarrow \left(\tilde{T}_i - \tilde{T}_{i+1}\right) + \left(\tilde{T}_i - \tilde{T}_{i-1}\right) \geq 0$$

Therefore $\tilde{T}_i$ for $i = 2, ...N-1$ is not the smallest component of $\tilde{\boldsymbol{T}}$. Now if the smallest component if $\tilde{T}_1$ then

$$-\tilde{T}_2 + 2\tilde{T}_1 \geq 0 \Longrightarrow \tilde{T}_1 \geq \tilde{T}_2 - \tilde{T}_1 \geq 0$$

Now if the smallest component if $\tilde{T}_N$ then

$$-\tilde{T}_{N-1} + 2\tilde{T}_N \geq 0 \Longrightarrow \tilde{T}_N \geq \tilde{T}_{N-1} - \tilde{T}_N \geq 0$$

Conclusion $\tilde{\boldsymbol{T}} \geq 0$ and $\underline{A}^{-1}$ is non negative.

Numerical scheme : properties and consequences

- Boundedness of $\underline{\mathcal{A}}^{-1}$: $0 \leq \sum_{j=1}^{N} \left(\underline{\mathcal{A}}^{-1}\right)_{ij} \leq \frac{1}{8}$ for $i = 1, \dots N$

Indeed the function $T(x) = \frac{x(1-x)}{2}$ is the solution both of the continuous and the discrete problem with respectively $S \equiv 1$ and $\boldsymbol{S} = \mathbf{1}$. Therefore (for this choise of $\boldsymbol{S}$)

$$0 \leq T(x_i) = \left(\underline{\mathcal{A}}^{-1}\mathbf{1}\right)_i = \sum_{j=1}^{N} \left(\underline{\mathcal{A}}^{-1}\right)_{ij} \leq \max_{\mathbf{x}\in(0,1)} \left(\frac{x(1-x)}{2}\right) = \frac{1}{8}$$

Discrete stability : $\tilde{\boldsymbol{T}} = \underline{\mathcal{A}}^{-1}\boldsymbol{S}$ and $\|\tilde{\boldsymbol{T}}\|_\infty = \leq \frac{1}{8}\|\boldsymbol{S}\|_\infty$

$$\|\tilde{\boldsymbol{T}}\|_\infty = \max_{i=1}^{N} |\tilde{T}_i| = \max_{i=1}^{N} \left|\left(\underline{\mathcal{A}}^{-1}\right)_{ij} S_j\right| \leq \max_{i=1}^{N} \left(\left|\left(\underline{\mathcal{A}}^{-1}\right)_{ij}\right| |S_j|\right)$$
$$\leq \left(\max_{i=1}^{N} \left|\left(\underline{\mathcal{A}}^{-1}\right)_{ij}\right|\right) \max_{j=1}^{N} |S_j| \leq \frac{1}{8}\|\boldsymbol{S}\|_\infty$$

Definitions for a Boundary Values Problem Scheme

Continuous $\mathcal{L}(\boldsymbol{T})$ and discrete $\tilde{\mathcal{L}}\left(\tilde{\boldsymbol{T}}\right)$ operators.

Examples : : $\mathcal{L}(\boldsymbol{T})_i = -\frac{d^2T}{d\mathbf{x}^2}(\mathbf{x}_i) = S_i$ and $\tilde{\mathcal{L}}\left(\tilde{\boldsymbol{T}}\right) = \underline{\mathcal{A}}\tilde{\boldsymbol{T}} = \boldsymbol{S}$

- Truncation error vector :

$$\boldsymbol{\mathcal{E}} = \tilde{\mathcal{L}}(\boldsymbol{T}) - \boldsymbol{S} - (\mathcal{L}(\boldsymbol{T}) - \boldsymbol{S}) = \tilde{\mathcal{L}}(\boldsymbol{T}) - \mathcal{L}(\boldsymbol{T}) = \tilde{\mathcal{L}}(\boldsymbol{T}) - \boldsymbol{S}$$

- discretization error vector : $\boldsymbol{e} = \boldsymbol{T} - \tilde{\boldsymbol{T}}$
- Error equation (linear case) :

$$\boldsymbol{\mathcal{E}} = \tilde{\mathcal{L}}(\boldsymbol{T}) - \boldsymbol{S} = \tilde{\mathcal{L}}(\boldsymbol{T}) - \tilde{\mathcal{L}}\left(\tilde{\boldsymbol{T}}\right) = \tilde{\mathcal{L}}\boldsymbol{e} \Longrightarrow \boldsymbol{e} = \tilde{\mathcal{L}}^{-1}\boldsymbol{\mathcal{E}}$$

- A-priory Error estimate : $\|\boldsymbol{e}\| \leq C\|\boldsymbol{\mathcal{E}}\|$

Definitions for a Boundary Values Problem Scheme

- Consistency and Accuracy : A scheme is consistent of p'th order accurate (with $p > 0$) if

$$|\mathcal{E}_i| = O(\delta \mathbf{x}^p) \quad \forall i \Longrightarrow \|\mathcal{E}\|_\infty = O(\delta \mathbf{x}^p) = C\delta \mathbf{x}^p$$

- L^∞ Stability : $\|\tilde{\mathcal{L}}^{-1}\|_\infty \leq C \quad \forall \delta \mathbf{x}$ where

$$\|\underline{\mathcal{A}}\|_\infty = \sup_{\boldsymbol{u} \in \mathbb{R}^N} \left(\frac{\|\underline{\mathcal{A}}\boldsymbol{u}\|_\infty}{\|\boldsymbol{u}\|_\infty} \right) = \sup_{\|\boldsymbol{u}\|_\infty = 1} \left(\|\underline{\mathcal{A}}\boldsymbol{u}\|_\infty \right) = \max_i \sum_{j=1}^{N} |\mathcal{A}_{ij}|$$

- L^p Stability : $\|\tilde{\mathcal{L}}^{-1}\|_p \leq C \quad \forall \delta \mathbf{x}$
- Convergence : $\lim_{\delta \mathbf{x} \to 0} \|\boldsymbol{e}\| = 0$

  Consistency + Stability = Convergence

$$\|\boldsymbol{e}\| = \|\tilde{\mathcal{L}}^{-1}\boldsymbol{\mathcal{E}}\| \leq \|\tilde{\mathcal{L}}^{-1}\|\|\mathcal{E}\| \leq C^* \delta \mathbf{x}^p$$

Application to $-\frac{d^2T}{d\mathbf{x}^2}(\mathbf{x}) = S(\mathbf{x})$, Scheme $\underline{\mathcal{A}}\tilde{T} = S$

$$R_2(\mathbf{x}, \beta\delta\mathbf{x}) = \beta^3 \frac{\delta\mathbf{x}^3}{6} \frac{d^3T}{d\mathbf{x}^3}(\mathbf{x}) + \beta^4 \frac{\delta\mathbf{x}^4}{24} \frac{d^4T}{d\mathbf{x}^4}(\xi_\beta)$$

- Consistency and Accuracy
  $\mathcal{E}(\mathbf{x}) = \frac{R_2(\mathbf{x},\delta\mathbf{x})+R_2(\mathbf{x},-\delta\mathbf{x})}{\delta\mathbf{x}^2} = \frac{\delta\mathbf{x}^2}{24}\left(\frac{d^4T}{d\mathbf{x}^4}(\xi_1) + \frac{d^4T}{d\mathbf{x}^4}(\xi_{-1})\right)$
  If T is $\mathcal{C}^4$-smooth then $\mathcal{E}(\mathbf{x}) = C\delta\mathbf{x}^2$.
  Then scheme is consistent of second order accurate.
- L^∞ Stability : we have proved that $\underline{\mathcal{A}}^{-1}$ is non negative (maximum principle) and bounded by $\frac{1}{8}$. Therefore

$$\|\tilde{\mathcal{L}}^{-1}\|_\infty = \sum_{j=1}^{N} \left|\left(\underline{\mathcal{A}}^{-1}\right)_{ij}\right| = \sum_{j=1}^{N} \left(\underline{\mathcal{A}}^{-1}\right)_{ij} \leq \frac{1}{8}$$

- A-priory Error estimate : $\|\boldsymbol{e}\|_\infty \leq \frac{1}{8}\|\boldsymbol{\mathcal{E}}\|_\infty$
- L^∞ Convergence : $\|\boldsymbol{e}\|_\infty \leq \frac{C}{8}\delta\mathbf{x}^2$ Then the scheme converge.

L^2 Convergence : Direct evaluation !

Eigenvalues λ_m and Eigenvectors $\boldsymbol{\vartheta}_m$ of $\underline{\mathcal{A}}$:

$$\begin{aligned} \lambda_m &= \frac{2 - 2\cos\theta_m}{\delta\mathbf{x}^2} = \frac{4}{\delta\mathbf{x}^2}\sin^2\frac{\theta_m}{2}, \qquad \text{with} \qquad \theta_m = m\pi\delta\mathbf{x} \\ \boldsymbol{\vartheta}_{m,j} &= \sin(j\theta_m) \end{aligned}$$

- Rayleigh Quotient : $\lambda_1 = \min_m \lambda_m \le \frac{(\boldsymbol{u}, \underline{\mathcal{A}}\boldsymbol{u})}{(\boldsymbol{u}, \boldsymbol{u})} \le \max_m \lambda_m = \lambda_N$
- Discrete L^2-norm consistent with the continuous L^2-norm :

$$\|\boldsymbol{u}\|_{\tilde{2}}^2 = \delta\mathbf{x}\, \boldsymbol{u} \cdot \boldsymbol{u}$$ caution : in the current case $N\delta\mathbf{x} \le 1$

- Error equation $\tilde{\mathcal{L}}\boldsymbol{e} = \boldsymbol{\mathcal{E}}$. By the vertue of the Cauchy-Schartz inequality and the consistency $|\mathcal{E}_i| = C\delta\mathbf{x}^2$ and Rayleigh quotient

$$(\boldsymbol{e}, \underline{\mathcal{A}}\boldsymbol{e}) = (\boldsymbol{e}, \boldsymbol{\mathcal{E}}) \le (\boldsymbol{e}\cdot\boldsymbol{e})^{\frac{1}{2}} (\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\mathcal{E}})^{\frac{1}{2}} \qquad \text{and } \lambda_1 (\boldsymbol{e}\cdot\boldsymbol{e}) \le (\boldsymbol{e}, \boldsymbol{\mathcal{E}})$$

$$\|\boldsymbol{e}\|_{\tilde{2}}^2 \le \frac{1}{\lambda_1}\|\boldsymbol{e}\|_{\tilde{2}}\|\boldsymbol{\mathcal{E}}\|_{\tilde{2}} \Longrightarrow \|\boldsymbol{e}\|_{\tilde{2}} \le \frac{1}{\lambda_1}\|\boldsymbol{\mathcal{E}}\|_{\tilde{2}} \le \frac{C}{\lambda_1}\delta\mathbf{x}^2$$

L^2 Solving the system and convergence

Principle of iterative method for solving $\underline{\mathcal{A}}\tilde{\boldsymbol{T}} = \boldsymbol{S}$

Define a convergent serie $\tilde{\boldsymbol{T}}^n$ such that $\tilde{\boldsymbol{T}} = \lim_{n\to\infty} \tilde{\boldsymbol{T}}^n$

- iteration error $\boldsymbol{e}^n = \tilde{\boldsymbol{T}} - \tilde{\boldsymbol{T}}^n$
- residual error : $\boldsymbol{r}^n = \boldsymbol{S} - \underline{\mathcal{A}}\tilde{\boldsymbol{T}}^n$
- error equation : $\underline{\mathcal{A}}\boldsymbol{e}^n = \boldsymbol{r}^n$

$$\kappa(\underline{\mathcal{A}}) = \frac{\lambda_N}{\lambda_1} = \frac{\sin^2 \frac{N\pi\delta\mathbf{x}}{2}}{\sin^2 \frac{\pi\delta\mathbf{x}}{2}},$$

$\underline{\mathcal{A}}$ is an ill-conditioned matrix on fine mesh.

$$\lim_{\delta\mathbf{x}\longrightarrow 0} \kappa(\underline{\mathcal{A}}) = \infty$$

Solving the system : Simple iteratives methods

Split of $\underline{\mathcal{A}}$ as Diagonal, Lower triangular, Upper triangular matrices.

$$\underline{\mathcal{A}} = \underline{D} - \underline{L} - \underline{U},$$

Error iteration : $\boldsymbol{e}^{n+1} = \underline{\mathcal{R}}\boldsymbol{e}^{n}$

- Jacobi : $\underline{\mathcal{R}} = \underline{\mathcal{R}}_J = \underline{D}^{-1}\left(\underline{L} + \underline{U}\right) = I - \frac{\delta \mathbf{x}^2}{2}\underline{\mathcal{A}}$

$$\tilde{\boldsymbol{T}}^{n+1} = \underline{D}^{-1}\left(\underline{L} + \underline{U}\right)\tilde{\boldsymbol{T}}^{n} + \underline{D}^{-1}\boldsymbol{S} = \tilde{\boldsymbol{T}}^{n} + \underline{D}^{-1}\boldsymbol{r}^{n}$$

- Gauss-Seidel : $\underline{\mathcal{R}} = \underline{\mathcal{R}}_{GS} = \left(\underline{D} - \underline{L}\right)^{-1}\underline{U}$

$$\tilde{\boldsymbol{T}}^{n+1} = \left(\underline{D} - \underline{L}\right)^{-1}\left(\underline{U}\tilde{\boldsymbol{T}}^{n} + \boldsymbol{S}\right) = \tilde{\boldsymbol{T}}^{n} + \left(\underline{D} - \underline{L}\right)^{-1}\boldsymbol{r}^{n}$$

- Conjugated-Gradient : (see Optimisation Lectures).

Congergence rate of Jacobi method (see Linear Algebra)

let us defined the initial error with the eigenvectors ϑ_m of the matrix $\underline{\mathcal{A}}$.

$$\boldsymbol{e}^0 = \sum_m \alpha_m \vartheta_m \quad \text{recall that} \quad \lambda_m = \frac{2}{\delta \mathbf{x}^2}\left(1 - \cos(m\pi\delta\mathbf{x})\right)$$

$$\underline{\mathcal{R}}_J \vartheta_m = \left(1 - \frac{\lambda_m \delta \mathbf{x}^2}{2}\right) \vartheta_m = \cos(m\pi\delta\mathbf{x})$$

Therefore

$$\boldsymbol{e}^n = \sum_m \alpha_m \left(\cos(m\pi\delta\mathbf{x})\right)^n \vartheta_m$$

Rate of convergence associated to the mode m is obtain by choosing $\boldsymbol{e}^0 = \vartheta_m$. In this case

$$\frac{\|\boldsymbol{e}^n\|}{\|\boldsymbol{e}^0\|} = |\cos(m\pi\delta\mathbf{x})|^n = \left|1 - \frac{(m\pi\delta\mathbf{x}^2)}{2} + ...\right|^n < 1$$

Plan

Plan

Plan

Plan

Plan