FEM and sparse linear system solving

Lecture 1, Sept 22, 2017: Introduction

http://people.inf.ethz.ch/arbenz/FEM17

Peter Arbenz
Computer Science Department, ETH Zürich
E-mail: arbenz@inf.ethz.ch

Introduction: Survey on lecture

1. The finite element method
2. Direct solvers for sparse systems
3. Iterative solvers for sparse systems

Introduction: Extended survey on lecture

- The finite element method
  - Introduction, model problems.
  - 1D problems. Piecewise polynomials in 1D.
  - 2D problems. Triangulations. Piecewise polynomials in 2D.
  - Variational formulations. Galerkin finite element method.
  - Implementation aspects.
- Direct solvers for sparse systems
  - LU and Cholesky decomposition
  - Sparse matrices, storage schemes.
  - Fill-reducing orderings.

Introduction: Extended survey on lecture (cont.)

- Iterative solvers for sparse systems
  - Stationary iterative methods, preconditioning.
  - Preconditioned conjugate gradient algorithm (PCG).
  - Krylov space methods for nonsymmetric systems
    Generalized Minimal RESidual (GMRES) algorithm.
  - Incomplete factorization preconditioning.
  - Multigrid preconditioning.
  - Indefinite problems (SYMMLQ, MINRES).
  - Nonsymmetric Lanczos iteration based methods
    Bi-CG, QMR, CGS, BiCGstab.

Literature

- M. G. Larson, F. Bengzon: *The Finite Element Method: Theory, Implementation, and Applications.* Springer, 2013.
- H. Elman, D. Sylvester, A. Wathen. *Finite elements and fast iterative solvers (2nd ed.).* Oxford University Press, 2014.
- Y. Saad. *Iterative methods for sparse linear systems (2nd ed.).* SIAM, 2003. (`www-users.cs.umn.edu/~saad/books.html`)
- T. Davis. *Direct Methods for Sparse Linear Systems.* SIAM, 2006.
- V. Dolean, P. Jolivet, F. Nataf. *An Introduction to Domain Decomposition Methods.* SIAM, 2015.
- F. Hecht. *FreeFem++* `http://www.freefem.org/`

Organization

- 13 lectures
- complementary exercises
  - to get hands-on experience
  - based on MATLAB and FreeFem++
    - MATLAB's finite element toolbox ($\leq$ 2D)
    - FEM used to construct matrices
    - Direct / iterative methods uses to solve the resulting systems of equations
- Examination
  - 30' oral examination
  - in exam session

The Poisson problem I

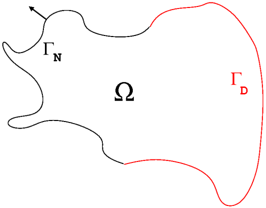
Model problem posed on a simply connected domain Ω:

$$
\begin{aligned}
-\Delta u(\mathbf{x}) := -\sum_{i=1}^{d} \frac{\partial^2 u(\mathbf{x})}{\partial x_i^2} &= f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \\
u(\mathbf{x}) &= g_D(\mathbf{x}), \quad \mathbf{x} \in \Gamma_D = \partial\Omega_D, \\
\frac{\partial u(\mathbf{x})}{\partial n} &= g_N(\mathbf{x}), \quad \mathbf{x} \in \Gamma_N = \partial\Omega_N.
\end{aligned}
\tag{M}
$$

Here, $\partial\Omega = \Gamma_D \cup \Gamma_N$. On $\Gamma_D \neq \emptyset$ *Dirichlet* boundary conditions are imposed. On Γ_N *Neumann* boundary conditions are imposed. *Mixed* or *Cauchy* or *Robin* boundary conditions are possible:

$$\alpha u + \beta \frac{\partial u}{\partial n} = g.$$

The Poisson problem II



$$\frac{\partial u}{\partial n} = \mathbf{grad}\ u \cdot \mathbf{n}$$

If we assume that $\Gamma_D \neq \emptyset$ then (M) has a unique solution.
REMARK: If $\Gamma_D = \emptyset$ then

$$-\int_{\partial\Omega} g_N \, d\mathbf{s} = -\int_{\partial\Omega} \frac{\partial u}{\partial n} \, d\mathbf{s} = -\int_{\Omega} \Delta u \, d\mathbf{x} = \int_{\Omega} f \, d\mathbf{x}.$$

is necessary for the existence of a solution.

Origins of the Poisson equation

- Steady state heat conduction (temperature distribution in homogeneous medium due to heat sources.
- Electrostatics: Electric potential due to charge distribution.
- Astronomy: Gravitational potential due to mass distribution.
- Fluid dynamics: Potential of irrotational (curl-free) flow.
- Deviation from equilibrium state of a membrane due to external forces.

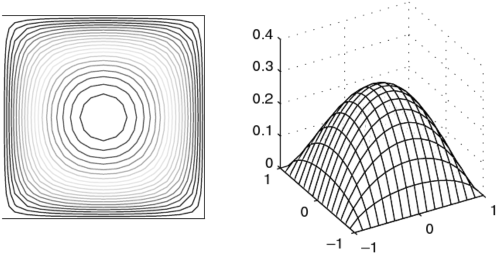
Origins of the Poisson equation (cont.)

- Previous problems assumed homogeneous medium.
- Non-homogeneity due to (e.g.) varying material properties can be taken into account and do not affect the ideas of the following analyses, except that they make things more complicated to write down.

$$
\begin{aligned}
-\operatorname{div}\left(K(\mathbf{x})\,\mathbf{grad}\,u(\mathbf{x})\right) &= f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad K(\mathbf{x}) \text{ positive definite}, \\
u(\mathbf{x}) &= g(\mathbf{x}), \quad \mathbf{x} \in \Gamma_D, \\
\frac{\partial (Ku(\mathbf{x}))}{\partial n} &= g_N(\mathbf{x}), \quad \mathbf{x} \in \Gamma_N.
\end{aligned}
$$

Example 1

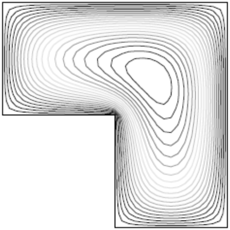
Solution for $\Omega = (-1, 1)^2$, right hand side $f = 1$, $g = 0$ where $\Gamma_D = \partial\Omega$.

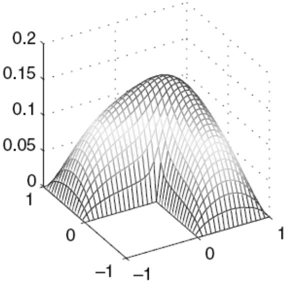


The solution is continuous up to the boundary and has second derivatives in the interior.
Analytic solution can be obtained by separation of variables.

Example 2 (same problem / different domain)

Solution for $\Omega =$ *L*-shaped membrane, right hand side $f = 1$, $g = 0$ where $\Gamma_D = \partial\Omega$.





The solution is continuous up to the boundary, but its first derivative is singular at the reentrant corner. There it has the form $u(r, \vartheta) = r^{2/3} \sin(2\vartheta + \pi)/3)$

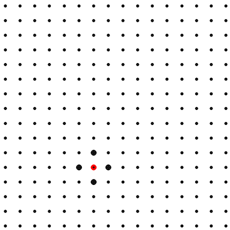
Finite differences

The Poisson equation $-\Delta u = f$ (plus some boundary conditions) on a domain with axis-parallel boundary is often solved by finite differences on a rectangular grid.

We define a rectangular grid with grid points that are a distance h apart. In each grid point the Laplacian of u can be expanded as

$$
\begin{aligned}
-\Delta u(x_1, x_2) &= -\frac{\partial^2}{\partial x_1^2} u(x_1, x_2) - \frac{\partial^2}{\partial x_2^2} u(x_1, x_2) \\
&\approx \frac{1}{h^2} \left(4u(x_1, x_2) - u(x_1 + h, x_2) - u(x_1 - h, x_2)\right. \\
&\qquad \left. -u(x_1, x_2 + h) - u(x_1, x_2 - h)\right) + \mathcal{O}(h^2)
\end{aligned}
$$

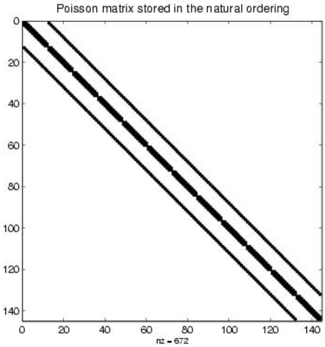
16×16 grid
grid width h
$\Omega = (0, 15h) \times (0, 15h)$



The discretization in every (interior) grid point is given by

$$4u_{\text{center}} - u_{\text{west}} - u_{\text{south}} - u_{\text{east}} - u_{\text{north}} = h^2 \cdot f_{\text{center}}$$

Structure of a FD matrix

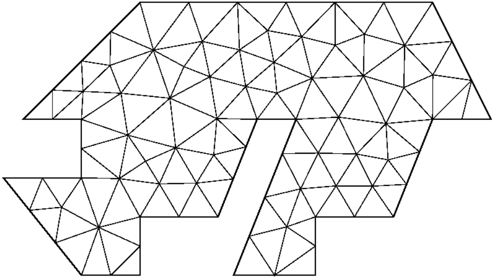


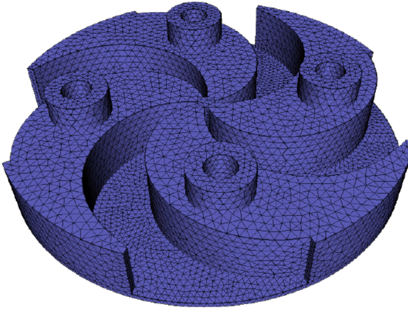
MATLAB spy of a matrix discretization of the Poisson equation $-\Delta u = f$ in $\Omega = (0,1)^2$, $u = 0$ on $\partial\Omega$, with finite differences on a 12×12 grid.

Conclusion on Finite Differences

- Finite differences are difficult (or cumbersome) to implement if the shapes of the domains get complicated, i.e., if the boundary is not aligned with the coordinate axes.
- *At this point finite elements come into play.*

Complicated domains





Structures that are difficult to discretize by finite differences.

The space of linear polynomials

$\mathbb{P}_1(I)$ vector space of linear polynomials on interval $I = [x_0, x_1]$

$$\mathbb{P}_1(I) = \{v : \; v(x) = c_0 + c_1 x, \; x \in I, \; c_0, c_1 \in \mathbb{R}\}.$$

$\{1, x\}$: the monomial basis for $\mathbb{P}_1(I)$

Definition: $\{x_0, x_1\}$: nodal points/nodes

$$\{\varphi_0(x), \varphi_1(x)\} : \; \varphi_j(x_i) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$ nodal basis for $\mathbb{P}_1(I)$

$$\varphi_0(x) = \frac{x_1 - x}{x_1 - x_0}, \quad \varphi_1(x) = \frac{x - x_0}{x_1 - x_0}$$

Any function v in $\mathbb{P}_1(I)$ can be expressed as

$$v(x) = \alpha_0 \varphi_0(x) + \alpha_1 \varphi_1(x)$$

$\{\alpha_0, \alpha_1\}$: node values $\alpha_0 = v(x_0); \; \alpha_1 = v(x_1)$

The space of continuous piecewise linear polynomials

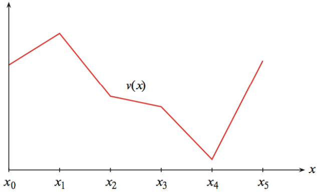
$\{x_i\}_{i=0}^n$: n+1 nodes in interval I:

$$0 = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = L.$$

Subintervals: $I_i = [x_{i-1}, x_i]$, $i = 1, \ldots, n$. Length: $h_i = x_i - x_{i-1}$.
The space of continuous piecewise linear functions V_h

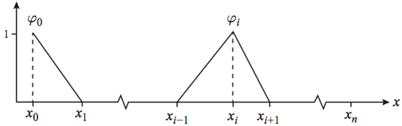
$$V_h = \left\{ v : v(x) \in C^0(I),\ v|_{I_i} \in \mathbb{P}_1(I_i) \right\}.$$

$C^0(I)$: the space of continuous functions on I
$\mathbb{P}_1(I_i)$: the space of linear functions on I_i



Any function v in V_h can be written as a linear combination of
$\{\varphi_i\}_{i=0}^n$: nodal bases (hat functions)
$\{\alpha_i\}_{i=0}^n$: nodal values of v.

$$v(x) = \sum_{i=0}^{n} \alpha_i \, \varphi_i(x) = \sum_{i=0}^{n} v(x_i) \, \varphi_i(x).$$



$$\varphi_i(x) = \begin{cases} (x - x_{i-1})/h_i, & \text{if } x \in I_i, \\ (x_{i+1} - x)/h_{i+1}, & \text{if } x \in I_{i+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Piecewise polynomial approximation

We consider two types of approximations of functions:

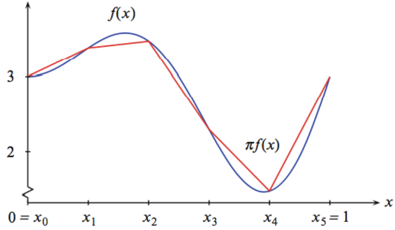
1. Piecewise (linear) interpolation.
   - Here we assume that the function f to be approximated (interpolated) is continuous.
   - What can happen if we try tro interpolate a function that is not continuous? (See Exercise 1.)
2. Piecewise (linear) best approximation in the least squares or L_2-sense.
   - Here we assume that the function f to be approximated is in $L_2(I)$, i.e.

$$\|v\|_{L^2(I)} \equiv \left(\int_I v^2 dx \right)^{1/2} < \infty.$$

Piecewise polynomial approximation technique I

1 Interpolation: Given a continuous function f on $I = [0, L]$, its continuous piecewise linear interpolant πf is defined by

$$\pi f(x) = \sum_{i=0}^{n} f(x_i)\varphi_i(x)$$



The interpolant πf approximates f exactly at the nodes x_i.

Interpolation error $f - \pi f$

There are many norms to measure the difference $f - \pi f$ in

1. the infinity norm
$$\|v\|_\infty = \max_{x \in I} |v(x)|$$
2. the $L^2(I)$ norm: for any square integrable function v on I
$$\|v\|_{L^2(I)} = \left(\int_I v^2 dx \right)^{1/2}$$

Note: $L^2(I)$ norm or any norm obeys the triangle inequality:

$$\|v + w\|_{L^2(I)} \leq \|v\|_{L^2(I)} + \|w\|_{L^2(I)}$$

Interpolation error $e = f - \pi f$ in $L^2(I)$-norm

Theorem 1: Linear interpolation in interval $I = [x_0, x_1]$ of length h.

$$\|f - \pi f\|_{L^2(I)} \leq C\, h^2 \|f''\|_{L^2(I)}$$
$$\|(f - \pi f)'\|_{L^2(I)} \leq C\, h \|f''\|_{L^2(I)}$$

Proof: Fundamental thm of calculus $\rightarrow$ $e(y) = e(x_0) + \int_{x_0}^{y} e' dx$

Using Cauchy-Schwarz inequality

$$e(y) = \int_{x_0}^{y} e' dx \leq \int_{x_0}^{y} |e'| dx \leq \left(\int_I 1 dx\right)^{\frac{1}{2}} \left(\int_I {e'}^2 dx\right)^{\frac{1}{2}} = h^{\frac{1}{2}} \left(\int_I {e'}^2 dx\right)^{\frac{1}{2}}$$

$$\|e\|_{L^2(I)} \leq h \|e'\|_{L^2(I)}, \text{ also } \|e'\|_{L^2(I)} \leq h \|e''\|_{L^2(I)} \Rightarrow \|e\|_{L^2(I)} \leq h^2 \|e''\|_{L^2(I)}$$

Interpolation error $e = f - \pi f$ in $L^2(I)$-norm (cont.)

Theorem 2: Piecewise linear interpolation in interval I.

$$\|f - \pi f\|^2_{L^2(I)} \leq C \sum_{i=1}^{n} h_i^4 \|f''\|^2_{L^2(I_i)}$$

$$\|(f - \pi f)'\|^2_{L^2(I)} \leq C \sum_{i=1}^{n} h_i^2 \|f''\|^2_{L^2(I_i)}$$

Proof: Note that

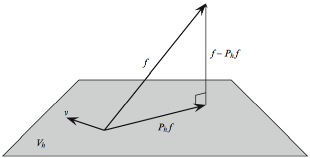
$$\|f - \pi f\|^2_{L^2(I)} = \sum_{i=1}^{n} \|f - \pi f\|^2_{L^2(I_i)}.$$

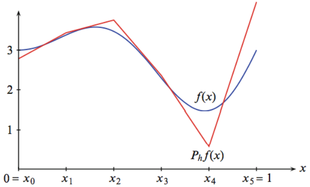
Then use triangle inequality and Theorem 1.

Piecewise polynomial approximation technique II

2 L^2 **-projection**: The L^2-projection $P_h f$ of f onto the space V_h is defined by

$$\int_I (f - P_h f) v_h \, dx = 0, \quad \forall v_h \in V_h \qquad (*)$$





The L^2-projection $P_h f$ approximates f on average. $P_h f$ commonly over- and undershoots local maxima and minima of f, respectively.

A priori error estimate

Theorem 3: The L_2-projection $P_h f$ satisfies the best approximation property:

$$\|f - P_h f\|_{L^2(I)} \leq \|f - v_h\|_{L^2(I)}, \quad \forall v_h \in V_h$$

Proof: Write $\|f - P_h f\|^2_{L^2(I)} = \int_I (f - P_h f)(f - v_h + v_h - P_h f)\, dx$, split equation, and use Schwarz inequality. □

Consequences:

$$\|f - P_h f\|^2_{L^2(I)} \leq \|f - \pi f\|^2_{L^2(I)}$$

$$\|f - P_h f\|^2_{L^2(I)} \leq C \sum_{i=1}^{n} h_i^4 \|f''\|^2_{L^2(I_i)}$$

Compute the L^2-projection

From $(*)$

$$\int_I (f - P_h f)\varphi_i \, dx = 0, \quad i = 0, 1, \ldots, n,$$

where φ_i are the hat functions. Since $P_h f$ belongs to V_h,

$$P_h f = \sum_{j=0}^{n} \xi_j \varphi_j,$$

where ξ_j are the unknown coefficients to be determined.

$$\int_I f \varphi_i dx = \int_I \left(\sum_{j=0}^{n} \xi_j \varphi_j \right) \varphi_i dx = \sum_{j=0}^{n} \xi_j \int_I \varphi_j \varphi_i dx, \quad i = 0, 1, \ldots, n.$$

Matrix form

$$\text{Mass matrix:}\quad m_{ij} = \int_I \varphi_j \varphi_i \, dx, \quad i,j = 0,1,\ldots,n.$$

$$\text{Load vector:}\quad b_i = \int_I f \varphi_i dx, \quad i = 0,1,\ldots,n.$$

$(n+1) \times (n+1)$ linear system for the $n+1$ unknown coefficients ξ_j

$$b_i = \sum_{j=0}^{n} m_{ij} \xi_j, \quad i = 0,1,\ldots,n.$$

Solve the linear system of equations $\mathbf{M}\xi = \mathbf{b}$.
Properties of **M**?

Quadrature rules

To compute the L_2-projection, compute integrals approximately

$$J = \int_I f(x)dx$$

- Step 1: Interpolate the integrand f by a polynomial
  - Midpoint rule: interpolation by polynomial of degree 0

    $$J \approx f(m)h \text{ where } m = (x_0 + x_1)/2$$

  - Trapezoidal rule: interpolation by polynomial of degree 1

    $$J \approx \frac{f(x_0) + f(x_1)}{2} h \text{ where } m = (x_0 + x_1)/2.$$

  - Simpson's formula: interpolation by polynomial of degree 2

    $$J \approx \frac{f(x_0) + 4f(m) + f(x_1)}{6} h \text{ where } m = (x_0+x_1)/2,\ h = x_1-x_0.$$

- Step 2: Integrate the interpolant.

Let's use Simpson's formula to integrate $m_{ij} = \int_I \varphi_j \varphi_i dx$

$$\varphi_i = \begin{cases} (x - x_{i-1})/h_i, & \text{if } x \in I_i \\ (x_{i+1} - x)/h_{i+1}, & \text{if } x \in I_{i+1} \\ 0, & \text{otherwise} \end{cases}$$

Entries of $\mathbf{M}$: $m_{ij} = 0$ for for $|i - j| > 1$.

$$m_{ii} = \int_I {\varphi_i}^2 dx = \int_{x_{i-1}}^{x_i} {\varphi_i}^2 dx + \int_{x_i}^{x_{i+1}} {\varphi_i}^2 dx$$

$$= \frac{0 + 4(\frac{1}{2})^2 + 1}{6} h_i + \frac{1 + 4(\frac{1}{2})^2 + 0}{6} h_{i+1} = \frac{h_i}{3} + \frac{h_{i+1}}{3},\ i = 1, 2, \cdots, n-1$$

First and last diagonal entries of $\mathbf{M}$ are $: \frac{h_1}{3}, \frac{h_n}{3}$ respectively

$$m_{ii+1} = \int_I \varphi_{i+1} \varphi_i dx = \int_{x_i}^{x_{i+1}} \varphi_{i+1} \varphi_i dx = \frac{h_{i+1}}{6}$$

Superdiagonal entries are same as subdiagonal entries.

$$\mathbf{M} = \begin{bmatrix} \frac{h_1}{3} & \frac{h_1}{6} & & & & \\ \frac{h_1}{6} & \frac{h_1}{3} + \frac{h_2}{3} & \frac{h_2}{6} & & & \\ & \frac{h_2}{6} & \frac{h_2}{3} + \frac{h_3}{3} & \frac{h_3}{6} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{h_{n-1}}{6} & \frac{h_{n-1}}{3} + \frac{h_n}{3} & \frac{h_n}{6} \\ & & & & \frac{h_n}{6} & \frac{h_n}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{h_1}{3} & \frac{h_1}{6} & & \\ \frac{h_1}{6} & \frac{h_1}{3} & & \\ & & & \\ & & & \end{bmatrix} + \begin{bmatrix} & & & \\ & \frac{h_2}{3} & \frac{h_2}{6} & \\ & \frac{h_2}{6} & \frac{h_2}{3} & \\ & & & \end{bmatrix} + \cdots + \begin{bmatrix} & & & \\ & & & \\ & & \frac{h_n}{3} & \frac{h_n}{6} \\ & & \frac{h_n}{6} & \frac{h_n}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{h_1}{3} & \frac{h_1}{6} & & & \\ \frac{h_1}{6} & \frac{h_1}{3} & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} + \begin{bmatrix} & & & & \\ & \frac{h_2}{3} & \frac{h_2}{6} & & \\ & \frac{h_2}{6} & \frac{h_2}{3} & & \\ & & & & \\ & & & & \end{bmatrix} + \cdots + \begin{bmatrix} & & & & \\ & & & & \\ & & & & \\ & & & \frac{h_n}{3} & \frac{h_n}{6} \\ & & & \frac{h_n}{6} & \frac{h_n}{3} \end{bmatrix}$$

$$= \sum_{i=1}^{n} \mathbf{I}_i \frac{h_i}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{I}_i^*$$

Here, $\mathbf{I}_i \in \mathbb{R}^{n\times 2}$ is a Boolean (0-1) matrix. It maps the local to the global degrees of freedom:

$$(\mathbf{I}_i)_{k\ell} = \begin{cases} 1, & k \text{ is the global number of the local node } \ell \\ 0, & \text{otherwise.} \end{cases}$$

Using the Trapezoidal rule:

$$
\begin{aligned}
b_i &= \int_I f\varphi_i dx = \int_{x_{i-1}}^{x_{i+1}} f\varphi_i dx = \int_{x_{i-1}}^{x_i} f\varphi_i dx + \int_{x_i}^{x_{i+1}} f\varphi_i dx \\
&\approx (f(x_{i-1})\varphi_i(x_{i-1}) + f(x_i)\varphi_i(x_i))h_i/2 \\
&\qquad + (f(x_i)\varphi_i(x_i) + f(x_{i+1})\varphi_i(x_{i+1}))h_{i+1}/2 \\
&= f(x_i)(h_i + h_{i+1})/2
\end{aligned}
$$

$$
\mathbf{b} = \begin{bmatrix} f(x_0)h_1/2 \\ f(x_1)(h_1+h_2)/2 \\ f(x_2)(h_2+h_3)/2 \\ \vdots \\ f(x_{n-1})(h_{n-1}+h_n)/2 \\ f(x_n)(h_n)/2 \end{bmatrix} = \sum_{i=1}^{n} \mathbf{I}_i \frac{h_i}{2} \begin{bmatrix} f(x_{i-1}) \\ f(x_i) \end{bmatrix}
$$

Compute the L^2-projection: summary

1. Create mesh with n elements (subintervals) on interval I and define corresponding space of continuous piecewise linear functions V_h.
2. Compute the matrix $\mathbf{M}$ and the vector $\mathbf{b}$:
$$m_{ij} = \int_I \varphi_j \varphi_i dx \quad \text{and} \quad b_i = \int_I f \varphi_i dx.$$
3. Solve the linear system
$$\mathbf{M}\,\xi = \mathbf{b}$$
4. Set
$$P_h f = \sum_{j=0}^{n} \xi_j \varphi_j$$
which is the best approximation of the function f.

The finite element method in 1D

- The finite element method for a model problem
- Variational formulation
- Finite element approximation
- Derivation of a linear system of equations
- Basic algorithm to compute the finite element solution

The finite element method for a model problem

Example: A two-point boundary value problem: Find u that satisfies

$$-u''(x) = f(x), \quad x \in I = [0, L]$$

Boundary conditions at interval endpoints: $u(0) = u(L) = 0$.

Question: How do we find u ?

- If $f = 1 \Rightarrow u = x(L-x)/2$ (analytic solution).
- For general f: difficult or even impossible to find the analytic solution u.

Variational formulation

To derive the finite element method, we first rewrite the differential equation as a variational equation. Multiplying $f = -u''$ by a test function v and integrating by parts we get

$$\int_0^L fv\,dx = -\int_0^L u''vdx = \int_0^L u'v'dx - u'(L)v(L) + u'(0)v(0).$$

Requirement: v and v' square integrable on I. Space of square-integrable functions is denoted by $L^2(I) = \left\{u : I \to \mathbb{R} | \int_I |u(x)|^2 < \infty\right\}$.

$$V_0 = \left\{v : ||v||_{L^2(I)} < \infty, ||v'||_{L^2(I)} < \infty, v(0) = v(L) = 0\right\}.$$

Variational formulation: Find $u \in V_0$ such that

$$\int_I u'v'\,dx = \int_I fv\,dx, \ \forall v \in V_0.$$

Finite element approximation

Approximate u by a continuous piecewise linear function.

- Mesh on the interval I consisting of n subintervals
- Corresponding space V_h of all continuous piecewise linears. $V_{h,0}$: subspace of those functions in V_h that satisfy the homogeneous Dirichlet boundary conditions

$$V_{h,0} = \{v \in V_h : v(0) = v(L) = 0\}.$$

Find $u_h \in V_{h,0}$ such that

$$\int_I u_h' v' \, dx = \int_I fv \, dx, \quad \forall v \in V_{h,0}.$$

Derivation of a linear system of equations

φ_i, $i = 1, \ldots, n-1$, are the hat functions spanning $V_{h,0}$

$$\int_I u_h' \varphi_i' \, dx = \int_I f \varphi_i \, dx, \qquad i = 1, 2, \ldots, n-1.$$

Since u_h belongs to $V_{h,0}$,

$$u_h = \sum_{j=1}^{n-1} \xi_j \varphi_j.$$

Unknown coefficients: ξ_j, $j = 1, 2, \ldots, n-1$

$$\int_I f \varphi_i dx = \int_I (\sum_{j=1}^{n-1} \xi_j \varphi_j') \varphi_i' dx = \sum_{j=1}^{n-1} \xi_j \int_I \varphi_j' \varphi_i' dx, \quad i = 1, 2, \ldots, n-1$$

Matrix form

$$b_i = \int_I f\varphi_i; \quad a_{ij} = \int_I \varphi_j'\varphi_i'\,dx, \quad i,j = 1,2,\dots,n-1$$

$(n-1)\times(n-1)$ linear system for the n-1 unknown coefficients ξ_j

$$b_i = \sum_{j=1}^{n-1} a_{ij}\,\xi_j, \quad i = 1,2,\dots,n-1$$

$$\mathbf{A}\,\xi = \mathbf{b}.$$

$\mathbf{A}$: $(n-1)\times(n-1)$ matrix $\rightarrow$ stiffness matrix
$\mathbf{b}$: $(n-1)\times 1 \rightarrow$ load vector

Basic finite element algorithm

1. Create a mesh with n elements on the interval I and define the corresponding space of continuous pcw linear functions $V_{h,0}$.
2. Compute matrix $\mathbf{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ and vector $\mathbf{b} \in \mathbb{R}^{(n-1)}$:

$$b_i = \int_I f\varphi_i, \quad A_{ij} = \int_I \varphi_j' \varphi_i' dx, \quad i,j = 1,2,\ldots,n-1.$$

3. Solve linear system

$$\mathbf{A}\xi = \mathbf{b}$$

4. Set

$$u_h = \sum_{j=1}^{n-1} \xi_j \varphi_j$$

A priori error estimate: $e = u - u_h$

Theorem 4 (Galerkin orthogonality): The finite element approximation u_h, satisfies the orthogonality

$$\int_I (u' - u_h')\, v_h'\, dx = 0, \quad \forall v_h \in V_{h,0}$$

Theorem 5 (Best approximation property) The finite element solution u_h satisfies

$$\|u' - u_h'\|_{L^2(I)} \leq \|u' - v_h'\|_{L^2(I)} \quad \forall v_h \in V_{h,0}.$$

Proof: Write $\|u' - u_h'\|^2_{L^2(I)} = \int_I (u' - u_h')(u' - v_h' + v_h' - u_h')\, dx$, split equation, and use Schwarz inequality. □

A priori error estimate: $e = u - u_h$ (cont.)

Theorem 6 (A priori error estimate) The finite element solution u_h satisfies the estimate

$$\|(u - u_h)'\|^2_{L^2(I)} \le C \sum_{i=1}^{n} h_i^2 \|u''\|^2_{L^2(I_i)} \quad \forall v \in V_{h,0}$$

Proof: Use best approximation property with $v = \pi u$ and Theorem 2. □

Consequences: With $h = \max_i h_i$ we get

$$\|u' - u_h'\|_{L^2(I)} \le C\, h\, \|u''\|_{L^2(I)}.$$

A model problem with variable coefficients

Consider the model problem, find u such that

$$-(au')' = f, \; x \in I = [0, L],$$
$$au'(0) = \kappa_0(u(0) - g_0),$$
$$-au'(L) = \kappa_L(u(L) - g_L).$$

$a(x) > 0$, $\kappa_0 \geq 0$, $\kappa_L \geq 0$, g_0 and g_L are given.
Step 1: Start by rewriting the differential equation as a variational equation.

$$\int_0^L fv dx = \int_0^L -(au')' v dx$$

Use integration by parts and substitute boundary conditions,

$$\int_I au'v'dx + \kappa_L u(L)v(L) + \kappa_0 u(0)v(0) = \int_I fvdx + \kappa_L g_L v(L) + \kappa_0 g_0 v(0)$$

Step 2: To obtain the finite element approximation $u_h \in V_h$ replace the continuous space V with the discrete space of continuous piecewise linears V_h in the variational formulation.

Find $u_h \in V_h$ such that

$$\int_I a u_h' v' dx + \kappa_L u_h(L) v(L) + \kappa_0 u_h(0) v(0) = \int_I f v dx + \kappa_L g_L v(L) + \kappa_0 g_0 v(0)$$

for all $v \in V_h$.

Step 3: In order to compute the finite element approximation u_h, write it as the linear combination of $\{\varphi_i\}_{i=0}^{n}$, hat functions are basis of V_h. φ_0 and φ_n: half hats at end points $x = 0$ and $x = L$.

$$u_h = \sum_{j=0}^{n} \xi_j \varphi_j$$

and derive the linear system of equations.

Step 4: Insert

$$u_h = \sum_{j=0}^{n} \xi_j \varphi_j$$

into

$$\int_I a u_h' v' dx + \kappa_L u_h(L) v(L) + \kappa_0 u_h(0) v(0) = \int_I f v dx + \kappa_L g_L v(L) + \kappa_0 g_0 v(0)$$

and choose $v = \varphi_i$, $i = 0, \ldots, n$,

$$\varphi_i = \begin{cases} (x - x_{i-1})/h_i, & \text{if } x \in I_i \\ (x_{i+1} - x)/h_{i+1}, & \text{if } x \in I_{i+1} \\ 0, & \text{otherwise} \end{cases}$$

and you will get ...

Assembly of the Stiffness Matrix and Load Vector

$$(\mathbf{A} + \mathbf{R})\xi = \mathbf{b} + \mathbf{r}.$$

Entries of $(n+1) \times (n+1)$ matrices $\mathbf{A}$, $\mathbf{R}$ and of $(n+1)$ vectors $\mathbf{b}$ and $\mathbf{r}$ are:

$$a_{ij} = \int_I a\varphi_j'\varphi_i' dx$$

$$r_{ij} = \kappa_L \varphi_j(L)\varphi_i(L) + \kappa_0 \varphi_j(0)\varphi_i(0)$$

$$b_i = \int_I f\varphi_i dx$$

$$r_i = \kappa_L g_L \varphi_i(L) + \kappa_0 g_0 \varphi_i(0)$$

Entries of $\mathbf{A}$: for $|i-j|>1 : a_{ij}=0$.

$$a_{i,i}=\int_I a\varphi_i'^2 dx=\int_{x_{i-1}}^{x_i} a\varphi_i'^2 dx+\int_{x_i}^{x_{i+1}} a\varphi_i'^2 dx=\frac{a_i}{h_i}+\frac{a_{i+1}}{h_{i+1}}$$

First and last diagonal entries of $\mathbf{A}$ are $:\frac{a_1}{h_1},\frac{a_n}{h_n}$ respectively

$$a_{i,i+1}=\int_I a\varphi_{i+1}'\varphi_i' dx=\int_{x_i}^{x_{i+1}} a\varphi_{i+1}'\varphi_i' dxs=-\frac{a_{i+1}}{h_{i+1}}$$

Superdiagonal entries are same as subdiagonal entries: $\mathbf{A}=\mathbf{A}^T$.
Entries of $\mathbf{R}$: $r_{ij}=\kappa_L\varphi_j(L)\varphi_i(L)+\kappa_0\varphi_j(0)\varphi_i(0)$ are all zero except i=j=0 or i=j=n.
Entries of $\mathbf{b}+\mathbf{r}$: done exactly as shown in the L^2 projection, additional terms $r_1=\kappa_0 g_0\varphi_i(0)$ and $r_n=\kappa_L g_L\varphi_i(L)$.

$$\mathbf{A}+\mathbf{R} = \begin{bmatrix} \frac{a_1}{h_1} & -\frac{a_1}{h_1} & & & & \\ -\frac{a_1}{h_1} & \frac{a_1}{h_1}+\frac{a_2}{h_2} & -\frac{a_2}{h_2} & & & \\ & -\frac{a_2}{h_2} & \frac{a_2}{h_2}+\frac{a_3}{h_3} & -\frac{a_3}{h_3} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\frac{a_{n-1}}{h_{n-1}} & \frac{a_{n-1}}{h_{n-1}}+\frac{a_n}{h_n} & -\frac{a_n}{h_n} \\ & & & & -\frac{a_n}{h_n} & -\frac{a_n}{h_n} \end{bmatrix} + \begin{bmatrix} \kappa_0 & & \\ & \ddots & \\ & & \kappa_L \end{bmatrix}$$

$$\mathbf{b}+\mathbf{r} = \begin{bmatrix} f(x_0)h_1/2 \\ f(x_1)(h_1+h_2)/2 \\ f(x_2)(h_2+h_3)/2 \\ \vdots \\ f(x_{n-1})(h_{n-1}+h_n)/2 \\ f(x_n)(h_n)/2 \end{bmatrix} + \begin{bmatrix} \kappa_0 g_0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \kappa_L g_L \end{bmatrix}$$

The global stiffness matrix $\mathbf{A}+\mathbf{R}$ can be written as

$$\mathbf{A}+\mathbf{R} = \frac{a_1}{h_1}\begin{bmatrix} 1 & -1 & & \\ -1 & 1 & & \\ & & & \\ & & & \end{bmatrix} + \frac{a_2}{h_2}\begin{bmatrix} & & & \\ & 1 & -1 & \\ & -1 & 1 & \\ & & & \end{bmatrix} + \cdots + \frac{a_n}{h_n}\begin{bmatrix} & & & \\ & & & \\ & & 1 & -1 \\ & & -1 & 1 \end{bmatrix} + \begin{bmatrix} \kappa_0 g_0 & & \\ & \ddots & \\ & & \kappa_L g_L \end{bmatrix}$$

$$= \mathbf{A}^{I_1} + \mathbf{A}^{I_2} + \cdots + \mathbf{A}^{I_n} + \mathbf{R}$$

Exercise

http://people.inf.ethz.ch/arbenz/FEM17/exercises/ex1.pdf