Lecture 2: Finite element methods for nonlinear PDEs

Pr. Ismail Merabet

Univ. of K-M-Ouargla

February 22, 2025

Contents

Introduction

In the first semester, we have treated linear problems.
In this lecture, we will generalize this to consider nonlinear problems.
To discretize with finite elements, we convert linear PDEs into linear variational problems. It is not a surprise that we will convert nonlinear PDEs into nonlinear variational problems!

Some examples.

Linear PDE	Nonlinear PDE
Stokes	Navier-Stokes
Linear elasticity	Hyperelasticity
Maxwell	Magnetohydrodynamics
Schrödinger	Gross-Pitaevskii
Linearised gravity	Einstein field equations

Example (Bratu-Gelfand equation).

For $\lambda \in \mathbb{R}$ we consider:

$$\begin{cases} u''(x) + \lambda e^u = 0, \\ u(0) = u(1) = 0. \end{cases} \tag{1}$$

By inspection we take $V = H_0^1(0,1)$. We multiply by a test function $v \in V$ and integrate by parts we obtain:

Variational problem

$$\begin{cases} \text{find } u \in V \text{ such that} \\ \displaystyle\int_0^1 u'(x)v'(x)\mathrm{d}x - \int_0^1 \lambda e^u v \, \mathrm{d}x = 0 \text{ for all } v \in V \end{cases} \tag{2}$$

the solutions of (1) are given by:

$$u(x) = 2\ln\left(\frac{\cosh(\alpha)}{\cosh(\alpha(1-2x))}\right), \quad \cosh(\alpha) = \frac{4}{\sqrt{2\lambda}}\alpha \tag{3}$$

$$\text{number of solutions } = \begin{cases} 2 & \lambda \in (0, \lambda^\star) \\ 1 & \lambda \in \{0, \lambda^\star\} \\ 0 & \text{otherwise} \end{cases}$$

with

$$\lambda^\star = 8\left(\min_{x>0} \frac{x}{\cosh x}\right)^2 \approx 3.5138307$$

Our basic abstraction for linear problems was:

$$\begin{cases} \text{find } u \in V, \text{ such that} \\ a(u, v) = F(v), \quad \forall v \in V. \end{cases} \tag{4}$$

Our abstraction for nonlinear problems will be:

$$\begin{cases} \text{find } u \in V, \text{ such that} \\ G(u; v) = 0, \quad \forall v \in V, \end{cases} \tag{5}$$

where:

$$G : V \times V \to \mathbb{R}.$$

We use $G(u; v)$ to record that G is nonlinear in u but always linear in v.

It's often useful to reformulate the variational statement as an equation, as we did in the linear case.
Define $H : V \to V^*$ via

$$(H(u))(v) = \langle H(u), v \rangle = G(u; v).$$

Solutions of our nonlinear variational problem are exactly roots of H, i.e. we seek $u \in V$ such that

$$H(u) = 0$$

We consider the Galerkin approximation:

$$\begin{cases} \text{find } u_h \in V_h \text{ such that} \\ G(u_h; v_h) = 0, \quad \forall v_h \in V_h. \end{cases} \tag{6}$$

A primary goal of an analysis is bounding the error $\|u - u_h\|_V$.

- But which u, and which u_h ?
- How do we pair the continuous and discrete solutions across mesh levels?
- How do we know that the discrete problem supports the right number of solutions?
- How do we know there are no spurious solutions?

These are difficult questions; possible to address, but we will sidestep them.

First, let's recall Newton's method in $\mathbb{R}$ and $\mathbb{R}^N$. Suppose that we want to solve :

$$f(x) = 0$$

The Newton algorithm:

- Start with a given x_0.
- Solve : $f'(x_k)d_k = -f(x_k)$.
- update: $x_{k+1} = x_k + d_k$.
- Stop: if $f(x_k) = 0$.

Remarks

Invertibility: We require $f'(x_k)$ to be invertible at every iteration.
Poor global convergence: The initial guess matters. With poor initial guesses, Newton's method may diverge to infinity, or get stuck in a cycle.
Good local convergence: If f is smooth, the solution is isolated, and the guess close, Newton converges quadratically.

This geometric reasoning is hard to generalise to higher dimensions. Let's look at a derivation that does extend.
Consider the Taylor expansion of f around x_k :

$$f\left(x_k + d_k\right) = f\left(x_k\right) + f'\left(x_k\right) d_k + \mathcal{O}\left(d_k^2\right).$$

Linearise the model by ignoring higher-order terms:

$$f\left(x_k + d_k\right) \approx f\left(x_k\right) + f'\left(x_k\right) d_k$$

and find d_k such that $f\left(x_k + d_k\right) \approx 0$:

$$0 = f\left(x_k\right) + f'\left(x_k\right) d_k$$

This does extend to an $F \in C^1\left(\mathbb{R}^N; \mathbb{R}^N\right)$. Newton's method is to solve

$$DF\left(x_k\right) d_k = -F\left(x_k\right); \text{ update } x_{k+1} = x_k + d_k,$$

where DF is the Jacobian (Fréchet derivative) of F.

The generalisation of Newton's method to Banach spaces is called the Newton-Kantorovich algorithm.
Kantorovich's theorem (1948) is a triumph, fundamental both to nonlinear analysis and applied mathematics. It does not assume the existence of a solution: given certain conditions on the residual and initial guess, it proves the existence and local uniqueness of a solution.
With a good initial guess, and great cleverness, it is possible to devise computer-assisted proofs of the existence of solutions to infinite-dimensional nonlinear problems.

Theorem (Kantorovich (1948))

Let X and Y be two Banach spaces. Let Ω be an open subset of X, the set where the residual is defined. Let $H \in C^1(\Omega, Y)$ be the residual of our nonlinear problem, and let $x_0 \in \Omega$ be an initial guess such that the Fréchet derivative $H'(x_0)$ is invertible (hence $H'(x_0) \in L(X; Y)$ and $H'(x_0)^{-1} \in L(Y; X)$). Let $B(x_0, r)$ denote the open ball of radius r centred at x_0.

Assume that there exists a constant $r > 0$ such that

(1) $\overline{B(x_0, r)} \subset \Omega$,

(2) $\left\| H'(x_0)^{-1} H(x_0) \right\|_X \leq \frac{r}{2}$,

(3) For all $\tilde{x}, x \in B(x_0, r)$,

$$\left\| H'(x_0)^{-1} \left(H'(\tilde{x}) - H'(x) \right) \right\|_{L(X;X)} \leq \frac{1}{r} \|\tilde{x} - x\|_X$$

Theorem (Kantorovich (1948))

Then

(1) $H'(x) \in L(X;Y)$ is invertible at each $x \in B(x_0, r)$.

(2) The Newton sequence $(x_k)_{k=0}^{\infty}$ defined by

$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k)$$

is such that $x_k \in B(x_0, r)$ for all $k \geq 0$ and converges to a root $x^\star \in \overline{B(x_0, r)}$ of H.

(3) For each $k \geq 0$,

$$\|x^\star - x_k\|_X \leq \frac{r}{2^k}$$

(4) The root $x^\star$ is the locally unique, i.e. $x^\star$ is the only root of H in the ball $\overline{B(x_0, r)}$.

The Bratu-Gelfand equation

$$u''(x) + \lambda e^u = 0, \quad u(0) = 0 = u(1)$$

has variational formulation: $u \in H_0^1(0,1)$ such that

$$G(u; v) = -\int_0^1 u'(x)v'(x)\mathrm{d}x + \int_0^1 \lambda e^u v \, \mathrm{d}x = 0$$

for all $v \in H_0^1(0,1)$.

Unwinding the statement of the Newton step, the update d_k solves

$$G_u\left(u_k; v, d_k\right) = -G\left(u_k; v\right) \text{ for all } v \in V$$

Here

$$G(u; v) = -\int_0^1 u'(x)v'(x)\mathrm{d}x + \int_0^1 \lambda e^u v \; \mathrm{d}x$$

so the Newton equation becomes :

$$\begin{cases} \text{find } d_k \in H_0^1(0,1), \text{ such that :} \forall v \in H_0^1(0,1) \\ \int_0^1 d_k'(x)v'(x)\mathrm{d}x - \int_0^1 \lambda e^{u_k} d_k v \; \mathrm{d}x = \int_0^1 u_k'(x)v'(x)\mathrm{d}x - \int_0^1 \lambda e^{u_k} v \; \mathrm{d}x. \end{cases} \tag{7}$$

Many questions remain. Are the linear equations well-posed? In general they are not coercive. We need a more general theory of well-posedness.
We will never compute the exact d_k; we can only compute a finite-dimensional approximation. One therefore develops inexact Newton-Kantorovich methods.
How can we control the approximation error of d_k ? Adaptive discretisations.
How can the algorithm be globalised? Line searches, trust regions, continuation in parameter, continuation in mesh.