Center for Mathematics and Artificial Intelligence

PDE Constrained Optimization: Implementation Aspects

Harbir Antil

Center for Mathematics and Artificial Intelligence (CMAI) and
Department of Mathematical Sciences
George Mason University

May 02, 2022

George Mason University

Outline

PDE Constrained Optimization – Overview

Derivative Computation

Application of Abstract Approach

Finite Element Method

Matlab Code: Newton CG, L-BFGS, Semismooth Newton

PDE Constrained Optimization (PDECO) Overview

Literature: [Lions, 1971, Tröltzsch, 2010, Ito and Kunisch, 2008, Antil et al., 2018, Hinze et al., 2009, Ulbrich, 2011, Attouch et al., 2006, Ekeland and Témam, 1999]

PDE constrained optimization has 4-5 major components:

- **control** z
- **state** u
- **state equation** (i.e., PDE) which associates with every control z a state u.
- **cost functional** J which depends on z and u, to be minimized;
- constraints obeyed by z (**control constraints**) and u (**state constraints**);

In the most abstract form this problem can be written as

$$\begin{aligned}
&\min J(u,z) && \\
&\text{s.t. } e(u,z) = 0, && \text{PDE} \\
&z \in Z_{ad}, && \text{control constraints} \\
&u \in U_{ad}, && \text{state constraints}
\end{aligned}$$

Abstract PDECO Problem

$$
\begin{aligned}
&\min J(u,z) && \\
&\text{s.t. } e(u,z) = 0, && \text{PDE} \\
&z \in Z_{ad}, && \text{control constraints} \\
&u \in U_{ad}, && \text{state constraints}
\end{aligned}
$$

where u: states, z: controls,

- $(U, \|\cdot\|_U)$, $(Y, \|\cdot\|_Y)$ Banach spaces, $(Z, \|\cdot\|_Z)$ is a reflexive Banach space.
- $Z_{ad} \subset Z$ and $U_{ad} \subset U$ are nonempty, closed convex set.
- $J : U \times Z \to \mathbb{R}$, $e : U \times Z_{ad} \to Y$ are sufficiently smooth mappings.
- Can think $(U, \|\cdot\|_U) = (\mathbb{R}^n, \|\cdot\|_2)$, $(Y, \|\cdot\|_Y) = (\mathbb{R}^n, \|\cdot\|_2)$, $(Z, \|\cdot\|_Z) = (\mathbb{R}^m, \|\cdot\|_2)$.

For simplicity we shall focus on the control constrained case, i.e., $U_{ad} = U$.

Three different formulations

- **Full space form**

$$
\begin{aligned}
&\min J(u,z) && \\
&\text{s.t. } e(u,z) = 0, && \text{PDE} \\
&z \in Z_{ad}, && \text{control constraints}
\end{aligned}
$$

- Let $x = (u,z)$, $X = U \times Z$, $X_{ad} = U \times Z_{ad}$ then we arrive at the following abstract problem

$$
\begin{aligned}
&\min J(x) && \\
&\text{s.t. } e(x) = 0, && \text{PDE} \\
&x \in X_{ad}. && \text{constraints}
\end{aligned}
$$

- **Reduced form.** Let us consider a map

$$S : \mathcal{Z} \to U, \quad z \mapsto S(z) = u \text{ with } e(S(z),z) = 0.$$

Here $\mathcal{Z}$ is open such that $Z_{ad} \subset \mathcal{Z} \subset Z$. We assume that for $z \in \mathcal{Z}$ there is a unique $u = S(z) \in U$. Then we can write

$$\min_{z \in Z_{ad}} \mathcal{J}(z) := J(S(z),z)$$

Outline

PDE Constrained Optimization – Overview

Derivative Computation

Application of Abstract Approach

Finite Element Method

Matlab Code: Newton CG, L-BFGS, Semismooth Newton

Derivative Computation

Compute gradient and Hessian information for

$$\mathcal{J}(z) = J(S(z), z), \quad \text{where } S(z) = u \text{ solves } e(u, z) = 0$$

Assumption

- J and e are twice continuously differentiable,
- $e_u(u, z)$ is continuously invertible.

Use adjoint equation approach for derivative computation.

See, e.g., Chapter 1 in [Hinze et al., 2009], Chapter 1 in [Antil et al., 2018] or [Heinkenschloss, 2008].

Gradient Computation: Reduced Form

- Derivative

$$\langle \mathcal{J}'(z), h \rangle_{Z^*,Z} = \langle J_u(S(z), z), S'(z)h \rangle_{U^*,U} + \langle J_z(S(z), z), h \rangle_{Z^*,Z}$$

- Implicit function theorem applied to $e(S(z), z) = 0$ gives

$$\begin{aligned} e_u(S(z), z)\left(S'(z)h\right) + e_z(S(z), z)h = 0 \\ \Longrightarrow S'(z)h = -e_u^{-1}(S(z), z)e_z(S(z), z). \end{aligned}$$

- Derivative $(u = S(z))$

$$\begin{aligned} &\langle \mathcal{J}'(z), h \rangle_{Z^*,Z} \\ &= \langle J_u(u, z), -e_u(u, z)^{-1} e_z(u, z)h \rangle_{U^*,U} + \langle J_z(u, z), h \rangle_{Z^*,Z} \\ &= \langle \underbrace{e_u(u, z)^{-*} J_u(u, z)}_{=:p}, -e_z(u, z)h \rangle_{Y^*,Y} + \langle J_z(u, z), h \rangle_{Z^*,Z} \\ &= \langle -e_z(u, z)^* p + J_z(u, z), h \rangle_{Z^*,Z} \end{aligned}$$

- Connection with Lagrangian $L(u,z,p) = J(u,z) - \langle p, e(u,z)\rangle_{Y^*,Y}$:
  - The adjoint variable p solves $e_u^*(u,z)p = J_u(u,z)$,
    which is equivalent to $L_u(u,z,p) = 0$.
  - Derivative

$$
\begin{aligned}
\langle \mathcal{J}'(z), h\rangle_{Z^*,Z} &= \langle -e_z(u,z)^* p + J_z(u,z), h\rangle_{Z^*,Z} \\
&= \langle L_z(u,z,p), h\rangle_{Z^*,Z}
\end{aligned}
$$

- Connection with Lagrangian $L(u,z,p) = J(u,z) - \langle p, e(u,z)\rangle_{Y^*,Y}$:
  - The adjoint variable p solves $e_u^*(u,z)p = J_u(u,z)$,
    which is equivalent to $L_u(u,z,p) = 0$.
  - Derivative

$$\begin{aligned}\langle \mathcal{J}'(z), h\rangle_{Z^*,Z} &= \langle -e_z(u,z)^* p + J_z(u,z), h\rangle_{Z^*,Z} \\ &= \langle L_z(u,z,p), h\rangle_{Z^*,Z}\end{aligned}$$

- Assuming that Z is a Hilbert space the **gradient** $\nabla\mathcal{J}(z)$ is such that

$$\begin{aligned}(\nabla\mathcal{J}(z), h)_Z &= \langle \mathcal{J}'(z), h\rangle_{Z^*,Z} \\ &= \langle -e_z(u,z)^* p + J_z(u,z), h\rangle_{Z^*,Z} \quad \forall h \in Z\end{aligned}$$

(Riesz representation Theorem)

- Connection with Lagrangian $L(u,z,p) = J(u,z) - \langle p, e(u,z)\rangle_{Y^*,Y}$:
  - The adjoint variable p solves $e_u^*(u,z)p = J_u(u,z)$,
    which is equivalent to $L_u(u,z,p) = 0$.
  - Derivative

$$
\begin{aligned}
\langle \mathcal{J}'(z), h\rangle_{Z^*,Z} &= \langle -e_z(u,z)^* p + J_z(u,z), h\rangle_{Z^*,Z} \\
&= \langle L_z(u,z,p), h\rangle_{Z^*,Z}
\end{aligned}
$$

- Assuming that Z is a Hilbert space the **gradient** $\nabla\mathcal{J}(z)$ is such that

$$
\begin{aligned}
(\nabla\mathcal{J}(z), h)_Z &= \langle \mathcal{J}'(z), h\rangle_{Z^*,Z} \\
&= \langle -e_z(u,z)^* p + J_z(u,z), h\rangle_{Z^*,Z} \quad \forall h \in Z
\end{aligned}
$$

  (Riesz representation Theorem)

Derivative Computation Using Adjoints

1. Given z, solve $e(u,z) = 0$ for u (if not done already).
2. Solve the adjoint equation $e_u(u(z),z)^* p = J_u(u(z),z)$ for p . Denote the solution by $p(z)$.
3. Compute $\mathcal{J}'(z) = J_z(u(z),z) - e_z(u(z),z)^* p(z)$.

Two PDE solves (possibly nonlinear in step 1, linear PDE in step 2)

Hessian Computation

- Apply **implicit differentiation** to

$$\mathcal{J}'(z) = J_z(u(z), z) - e_z(u(z), z)^* p(z)$$

to compute Hessian information.

Hessian Computation

- Apply **implicit differentiation** to

$$\mathcal{J}'(z) = J_z(u(z), z) - e_z(u(z), z)^* p(z)$$

  to compute Hessian information.

- This leads to

$$\mathcal{J}''(z) = T(S(z), z)^* H(S(z), z, p) T(S(z), z)$$

  where

$$T(u, z) = \begin{pmatrix} -e_u(u, z)^{-1} e_z(u, z) \\ I_Z \end{pmatrix}$$

  and $H(u, z, p)$ is given by

$$H(u, z, p) = \begin{pmatrix} L_{uu}(u, z, p) & L_{uz}(u, z, p) \\ L_{zu}(u, z, p) & L_{zz}(u, z, p) \end{pmatrix}$$

Hessian Computation

- Apply **implicit differentiation** to
$$\mathcal{J}'(z) = J_z(u(z), z) - e_z(u(z), z)^* p(z)$$
to compute Hessian information.
- This leads to
$$\mathcal{J}''(z) = T(S(z), z)^* H(S(z), z, p) T(S(z), z)$$
where
$$T(u, z) = \begin{pmatrix} -e_u(u, z)^{-1} e_z(u, z) \\ I_Z \end{pmatrix}$$
and $H(u, z, p)$ is given by
$$H(u, z, p) = \begin{pmatrix} L_{uu}(u, z, p) & L_{uz}(u, z, p) \\ L_{zu}(u, z, p) & L_{zz}(u, z, p) \end{pmatrix}$$
- In practice we are interested in computing **Hessian times vector**. For instance in the absence of control constraints we need to solve
$$\mathcal{J}'(z) = 0$$
which we can do so by solving Newton's equation for s_z in an iterative fashion
$$\mathcal{J}''(z) s_z = -\nabla \mathcal{J}(z).$$

Hessian-Times-Vector Computation

- **Hessian-Times-Vector Computation**
  1. Given z, solve $e(u,z)=0$ for u (if not done already).
  2. Solve adjoint equation: $e_u(u,z)^* p = J_u(u,z)$ for p (if not done already).
  3. Solve $e_u(u,z)w = -e_z(u,z)v$.
  4. Solve $e_u(u,z)^* q = D_{uu}L(u,z,p)w + D_{uz}L(u,z,p)v$.
  5. Compute
  $$\mathcal{J}''(z)v = -e_z(u,z)^* q + L_{uz}(u,z,p)w + L_{zz}(u,z,p)v.$$

  Two linear PDE solves in steps 3+4 per direction v.

- Vector s_z solves Newton equation $\mathcal{J}''(z)s_z = -\mathcal{J}'(z)$ if and only if (s_u, s_z) solves the quadratic program

$$\begin{aligned}
\min \quad & \left\langle \begin{bmatrix} J_u(\cdot) \\ J_z(\cdot) \end{bmatrix}, \begin{bmatrix} s_u(\cdot) \\ s_z(\cdot) \end{bmatrix} \right\rangle + \frac{1}{2}\left\langle \begin{bmatrix} s_u(\cdot) \\ s_z(\cdot) \end{bmatrix}, \begin{bmatrix} L_{uu}(\cdot\cdot) & L_{uz}(\cdot\cdot) \\ L_{zu}(\cdot\cdot) & L_{zz}(\cdot\cdot) \end{bmatrix} \begin{bmatrix} s_u(\cdot) \\ s_z(\cdot) \end{bmatrix} \right\rangle \\
\text{s.t.} \quad & e_u(\cdot)s_u + e_z(\cdot)s_z = 0,
\end{aligned}$$

where $(\cdot) = (u(z), z)$ and $(\cdot\cdot) = (u(z), z, p(z))$.

Outline

Semilinear Elliptic PDECO Problem

- Problem:

$$\text{Minimize } J(u,z) = \frac{1}{2}\int_\Omega |u(x) - u_d(x)|^2 \,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega |z(x)|^2 \,\mathrm{d}x$$

s.t.

$$\begin{aligned} -\mathrm{div}(A\nabla u) + f(\cdot, u) &= z \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

$$z \in Z_{ad} := \{z \in L^\infty(\Omega) \ : \ z_a(x) \le z(x) \le z_b(x) \quad \text{a.e. in } \Omega\}$$

with $z_a, z_b \in L^\infty(\Omega)$ and $z_a < z_b$ a.e. in Ω.

- **Control space:** $L^\infty(\Omega)$. For existence we use $L^p(\Omega)$ with $N/2 < p < \infty$. Notice that $Z_{ad} \subset L^\infty(\Omega) \subset L^p(\Omega)$.

Semilinear Elliptic PDECO Problem

- Problem:

$$\text{Minimize } J(u,z) = \frac{1}{2}\int_\Omega |u(x) - u_d(x)|^2 \,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega |z(x)|^2 \,\mathrm{d}x$$

s.t.

$$\begin{aligned} -\mathrm{div}(A\nabla u) + f(\cdot, u) &= z \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

$$z \in Z_{ad} := \{z \in L^\infty(\Omega) \;:\; z_a(x) \le z(x) \le z_b(x) \quad \text{a.e. in } \Omega\}$$

with $z_a, z_b \in L^\infty(\Omega)$ and $z_a < z_b$ a.e. in Ω.

- **Control space:** $L^\infty(\Omega)$. For existence we use $L^p(\Omega)$ with $N/2 < p < \infty$. Notice that $Z_{ad} \subset L^\infty(\Omega) \subset L^p(\Omega)$.
  State space: $U = L^\infty(\Omega) \cap H_0^1(\Omega)$. [Antil et al., 2017, Antil and Warma, 2020].

Semilinear Elliptic PDECO Problem

- Problem:

$$\text{Minimize } J(u,z) = \frac{1}{2}\int_\Omega |u(x) - u_d(x)|^2 \,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega |z(x)|^2 \,\mathrm{d}x$$

s.t.

$$\begin{aligned} -\mathrm{div}(A\nabla u) + f(\cdot, u) &= z \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

$$z \in Z_{ad} := \{z \in L^\infty(\Omega) \;:\; z_a(x) \le z(x) \le z_b(x) \quad \text{a.e. in } \Omega\}$$

with $z_a, z_b \in L^\infty(\Omega)$ and $z_a < z_b$ a.e. in Ω.

- **Control space:** $L^\infty(\Omega)$. For existence we use $L^p(\Omega)$ with $N/2 < p < \infty$. Notice that $Z_{ad} \subset L^\infty(\Omega) \subset L^p(\Omega)$.
State space: $U = L^\infty(\Omega) \cap H_0^1(\Omega)$. [Antil et al., 2017, Antil and Warma, 2020].
- Notice that $Z \subset L^\infty(\Omega)$. This is a sufficient condition for the differentiability of S, a Nemytskii (superposition) operator.

- Lagrangian:

$$L(y,u,p) = \frac{1}{2}\int_\Omega |u(x)-u_d(x)|^2\,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega |z(x)|^2\,\mathrm{d}x \\ - \int_\Omega \big(A\nabla u\cdot\nabla p + f(x,u)p - zp\big)\,\mathrm{d}x.$$

- Adjoint equation:

$$\begin{aligned} -\mathrm{div}(A^T\nabla p) + f_u(\cdot,u)p &= (u-u_d) \quad \text{in } \Omega \\ p &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

- Derivative

$$\langle \mathcal{J}'(z), h\rangle_{Z^*,Z} = \langle \lambda z + p, h\rangle_{Z^*,Z}.$$

- Gradient: $\nabla\mathcal{J}$ is computed as follows:

$$(\nabla\mathcal{J}(z), h)_Z = \langle \mathcal{J}'(z), h\rangle_{Z^*,Z}, \quad \forall h\in Z.$$

- Hessian: $L_{uz} = L_{zu} = 0$. Then use the approach described before.

Linear Parabolic PDECO Problem

- Problem:

$$\text{Minimize } \frac{1}{2}\int_0^T \int_\Omega |u(x,t) - u_d(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{\lambda}{2}\int_0^T \int_\Omega |z(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t$$

where u solves

$$\begin{aligned}
\partial_t u(x,t) - \text{div }(A\nabla u) + \mathbf{b}\cdot\nabla u &= z \quad \text{in } \Omega \times (0,T) \\
u &= 0, \quad \text{on } \Gamma_D \times (0,T), \\
A\nabla u \cdot \boldsymbol{\nu} &= 0 \quad \text{on } \Gamma_N \times (0,T) \\
u(x,0) &= 0 \quad \text{in } \Omega.
\end{aligned}$$

- Lagrangian (formally)

$$\begin{aligned}
L(u,z,p) =& \frac{1}{2}\int_0^T \int_\Omega |u(x,t) - u_d(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{\lambda}{2}\int_\Omega |z(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t \\
& \int_0^T \int_\Omega \big(\partial_t u(x,t)p - \text{div }(A\nabla u)p + \mathbf{b}\cdot\nabla u p - zp\big) \,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

- Adjoint equation

$$
\begin{aligned}
-\partial_t p(x,t) - \text{div}\,(A^T \nabla p) - \text{div}(\mathbf{b}p) &= (u - u_d) \quad \text{in } \Omega \times (0,T) \\
p &= 0, \quad \text{on } \Gamma_D \times (0,T), \\
(A\nabla p + \mathbf{b}p) \cdot \boldsymbol{\nu} &= 0 \quad \text{on } \Gamma_N \times (0,T) \\
p(x,T) &= 0 \quad \text{in } \Omega.
\end{aligned}
$$

- Gradient

$$
(\nabla \mathcal{J}(z), h)_Z = \langle \mathcal{J}'(z), h \rangle_{Z^*,Z} = \langle p + \lambda z, h \rangle_{Z^*,Z} \quad h \in Z.
$$

where $Z = L^2(0,T; L^2(\Omega))$.

Outline

Linear elliptic PDE and Weak Form

Literature: [Braess, 1997, Ciarlet, 2002, Brenner and Scott, 2008, Ern and Guermond, 2004]

Poisson's equation. Let Ω be a bounded open Lipschitz domain with boundary $\partial\Omega$. Consider the following elliptic boundary value problem where given f we seek u solving

$$\begin{aligned} -\mathrm{div}(A\nabla u) &= f \quad \text{in } \Omega \\ u &= g_D \quad \text{on } \partial\Omega_D \\ A\nabla u \cdot \boldsymbol{\nu} &= g_N \quad \text{on } \partial\Omega_N . \end{aligned}$$

Linear elliptic PDE and Weak Form

Literature: [Braess, 1997, Ciarlet, 2002, Brenner and Scott, 2008, Ern and Guermond, 2004]

Poisson's equation. Let Ω be a bounded open Lipschitz domain with boundary $\partial\Omega$. Consider the following elliptic boundary value problem where given f we seek u solving

$$\begin{aligned} -\text{div}(A\nabla u) &= f \quad \text{in } \Omega \\ u &= g_D \quad \text{on } \partial\Omega_D \\ A\nabla u \cdot \boldsymbol{\nu} &= g_N \quad \text{on } \partial\Omega_N. \end{aligned}$$

Weak solution. We define

$$V := \{v \in H^1(\Omega) \,:\, v|_{\partial\Omega_D} = g_D\}, \quad V_0 := \{v \in H^1(\Omega) \,:\, v|_{\partial\Omega_D} = 0\}.$$

We seek $u \in V$ such that

$$\int_\Omega A\nabla u \cdot \nabla v dx = \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds, \quad \forall v \in V_0.$$

Weak form cont.

Assume that there exist $w \in H^1(\Omega)$ such that $w|_{\partial\Omega_D} = g_D$. Then we write $u = u_0 + w$ with $u_0|_{\partial\Omega_D} = 0$.

Thus we are reduced to finding $u_0 \in V_0$ solving:

$$\int_\Omega A\nabla u_0 \cdot \nabla v dx = \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds - \int_\Omega A\nabla w \cdot \nabla v dx, \quad \forall v \in V_0.$$

Then $u = u_0 + w$.

- **Weak form.** Find $u_0 \in V_0$ solving:

$$\int_\Omega A\nabla u_0 \cdot \nabla v dx = \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds - \int_\Omega A\nabla w \cdot \nabla v dx, \quad \forall v \in V_0.$$

- **Galerkin approximation.** Find subspace $V_h \subset V_0$ with basis $\{v_1, \ldots, v_n\}$. Approximate u_0 by $u_h = \sum_{i=1}^n u_j v_j$.

  Find $u_h \in V_h$ such that $u_h \in V_h$ solves

$$\int_\Omega A\nabla u_h \cdot \nabla v dx = \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds - \int_\Omega A\nabla w \cdot \nabla v dx, \quad \forall v \in V_h$$

  or equivalently

$$\int_\Omega A\nabla u_h \cdot \nabla v_i dx = \int_\Omega f v_i dx + \int_{\partial\Omega_N} g_N v_i ds - \int_\Omega A\nabla w \cdot \nabla v_i dx, \quad i = \overline{1, n}.$$

Galerkin Approximation

Find $u_h \in V_h$ solving:

$$\int_\Omega A\nabla u_h \cdot \nabla v_i dx = \int_\Omega f v_i dx + \int_{\partial\Omega_N} g_N v_i ds - \int_\Omega A\nabla w \cdot \nabla v_i dx, \quad i = \overline{1,n}.$$

Equivalently we can write a matrix vector system:

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

where $\mathbf{u} = (u_1, \ldots, u_n)^T$, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$ with entries

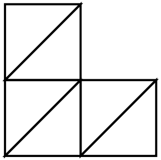
$$\begin{aligned}\mathbf{A}_{ij} &= \int_\Omega A\nabla v_j \cdot \nabla v_i \\ \mathbf{b}_i &= \int_\Omega f v_i dx + \int_{\partial\Omega_N} g_N v_i ds - \int_\Omega A\nabla w \cdot \nabla v_i dx, \quad i = \overline{1,n}.\end{aligned}$$

Piecewise linear basis functions on triangles

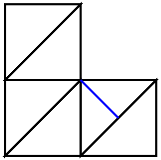
- We haven't specified V_h or basis functions $v_1, \ldots, v_n$ yet.
- We will define V_h as a subspace with piecewise polynomials.
- For simplicity let us consider $\Omega \subset \mathbb{R}^2$, and consider the subspace V_h of functions that are continuous on $\overline{\Omega}$ and piecewise linear basis functions on triangles.

Assume polygonal domain Ω and subdivide into open triangles $K_1, \ldots, K_{\text{elem}}$ such that

$$
\begin{aligned}
\overline{\Omega} &= \cup_i \overline{K}_i, \\
K_i \cap K_j &= \emptyset, \quad i \neq j, \\
\overline{K}_i \cap \overline{K}_j &= \emptyset, \text{ or } = \text{a vertex, or } = \text{ an edge of both } K_i \text{ and } K_j.
\end{aligned}
$$



Admissible triangulation



Inadmissible triangulation

- Subspace V_h consists of functions that are globally continuous on $\overline{\Omega}$ and piecewise linear on the triangles.
- Let $\mathbf{n}_1, \ldots, \mathbf{n}_{n_n} \in \overline{\Omega}$ are the nodes (same as vertices for piecewise linear elements). Here n_n is the number of nodes in the mesh.
- Basis functions $N_1, \ldots, N_{n_n}$ are piecewise linear on triangles and satisfy

$$N_i(\mathbf{n}_j) = \delta_{ij}.$$

  (refer to as *nodal basis*).
- Notice that $N_i(x)$ is nonzero only if x is contained in a element K that contains the node $\mathbf{n}_i$, i.e.,

$$N_i(x) = 0 \quad \text{for } x \notin \cup_{\mathbf{n}_i \in \overline{K}} \overline{K}.$$

- For notation simplicity we assume that $\mathbf{n}_{n+1}, \ldots, \mathbf{n}_{n_n}$ are the nodes corresponding to the Dirichlet boundary conditions, i.e.,
  $\mathbf{n}_1, \ldots \mathbf{n}_n \notin \partial\Omega_D$,
  $\mathbf{n}_{n+1}, \ldots \mathbf{n}_{n_n} \in \partial\Omega_D$.
- The subspace V_h has basis $\{N_1, \ldots, N_n\}$.
- Thus the Galerkin approximation now reads

$$\mathbf{A}\mathbf{u} = \mathbf{b},$$

where

$$\begin{aligned}
\mathbf{A}_{ij} &= \int_\Omega A\nabla N_j \cdot \nabla N_i \\
\mathbf{b}_i &= \int_\Omega f N_i dx + \int_{\partial\Omega_N} g_N N_i ds - \int_\Omega A\nabla w \cdot \nabla N_i dx, \quad i = \overline{1, n}.
\end{aligned}$$

- We will incorporate the Dirichlet boundary conditions later.
- Let us first compute $\mathbf{A} \in \mathbb{R}^{n_n,n_n}$ with entries

$$\mathbf{A}_{ij} = \int_\Omega A\nabla N_j \cdot \nabla N_i$$
$$\mathbf{b}_i = \int_\Omega f N_i dx + \int_{\partial\Omega_N} g_N N_i ds.$$

- We notice that

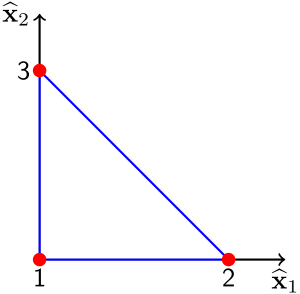
$$\int_\Omega A\nabla N_j \cdot \nabla N_i = \sum_{\ell=1}^{n_{\text{elem}}} \int_{K_\ell} A\nabla N_j \cdot \nabla N_i$$

and

$$\int_{K_\ell} A\nabla N_j \cdot \nabla N_i = 0$$

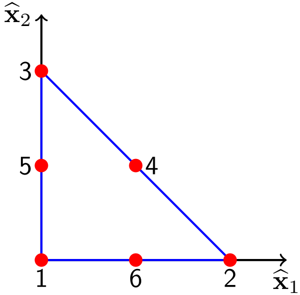
if node $\mathbf{n}_i$ or $\mathbf{n}_j$ is not in $\overline{K}_\ell$.

- For piecewise linear basis functions there are $n_{\text{basis}} = 3$ nodes per element. For each element K_ℓ compute $n_{\text{basis}} \times n_{\text{basis}}$ integrals. The resulting matrices are called *local*.
- Compute these integrals by transforming onto a reference element $\widehat{K}$ with vertices $(0,0),(1,0),(1,1)$.
- The linear basis on the reference triangle are given by



$$
\begin{aligned}
\widehat{N}_1(\widehat{x}) &= 1 - \widehat{x}_1 - \widehat{x}_2 \\
\widehat{N}_2(\widehat{x}) &= \widehat{x}_1 \\
\widehat{N}_3(\widehat{x}) &= \widehat{x}_2
\end{aligned}
$$

Quadratic Basis Functions on Reference Triangle



$$
\begin{aligned}
\widehat{N}_1(\widehat{x}) &= (1-\widehat{x}_1-\widehat{x}_2)(1-2\widehat{x}_1-2\widehat{x}_2)\\
\widehat{N}_2(\widehat{x}) &= \widehat{x}_1(2\widehat{x}_1-1)\\
\widehat{N}_3(\widehat{x}) &= \widehat{x}_2(2\widehat{x}_2-1)\\
\widehat{N}_4(\widehat{x}) &= 4\widehat{x}_1\widehat{x}_2\\
\widehat{N}_5(\widehat{x}) &= 4\widehat{x}_2(1-\widehat{x}_1-\widehat{x}_2)\\
\widehat{N}_6(\widehat{x}) &= 4\widehat{x}_1(1-\widehat{x}_1-\widehat{x}_2)
\end{aligned}
$$

- **Element transformation.** The map

$$
\begin{aligned}
F_E : \widehat{K} &\to K, \\
\widehat{x} &\mapsto F_K(x) = \mathbf{v}_{k_1} + B\widehat{x}
\end{aligned}
$$

  where $B = (\mathbf{v}_{k_2} - \mathbf{v}_{k_1}, \mathbf{v}_{k_3} - \mathbf{v}_{k_1}) \in \mathbb{R}^{2\times 2}$ is bijective and maps vertex $\widehat{\mathbf{v}}_i$ to $\mathbf{v}_{k_i}$, $i = 1, 2, 3$.
- The basis function on element K is given by

$$
N_{k_i}(x) = \widehat{N}_i(F_K^{-1}(x)), \quad i = 1, \ldots, n_{\text{basis}}.
$$

- Derivative. $\nabla N_{k_i}(x) = B^{-T} \nabla \widehat{N}_i(\widehat{x})$.

Assembly

- Let K be an element with nodes $\mathbf{n}_{k_1}, \ldots, \mathbf{n}_{k_{n_{\text{basis}}}}$.
- The integral
$$\int_K A(x) \nabla N_j(x) \cdot \nabla N_i(x)\, dx$$
is only nonzero if $i, j \in \{k_1, \ldots, k_{n_{\text{basis}}}\}$.
- If $i = k_\ell$ and $j = k_m$ then

$$\begin{aligned}\int_K A(x) \nabla N_j(x) \cdot \nabla N_i(x)\, dx &= \int_K A(x) \nabla N_{k_m}(x) \cdot \nabla N_{k_\ell}(x)\, dx \\ &= |\mathsf{det}(B)| \int_{\widehat{K}} A(F_K(\widehat{x})) \Big(B^{-T} \nabla \widehat{N}_m\Big) \Big(B^{-T} \nabla \widehat{N}_\ell\Big)\end{aligned}$$

Dirichlet Boundary Conditions

- Let n_n be the number of nodes in the mesh.
- $\mathbf{n}_{n+1}, \ldots, \mathbf{n}_{n_n}$ be the Dirichlet nodes. The subspace V_h has basis $\{N_1, \ldots, N_n\}$.
- So far we have assembled $\mathbf{A} \in \mathbb{R}^{n_n \times n_n}$ and $\mathbf{b} \in \mathbb{R}^{n_n}$.
- Approximate w by

$$w(x) \approx \sum_{j=n+1}^{n_n} d(\mathbf{n}_j) N_j(x)$$

  (Interpolation at nodes).
- We then modify $\mathbf{b}$ to incorporate nonzero boundary conditions as

$$\mathbf{b}_i = \int_\Omega f N_i + \int_{\partial\Omega} g_N N_i ds - \sum_{j=n+1}^{n_n} \mathbf{A}_{ij} d(\mathbf{n}_j).$$

- Then we need to solve

$$\mathbf{A}(1:n, 1:n)\mathbf{u}(1:n) = \mathbf{b}(1:n),$$

  where $1:n$ corresponds to the free nodes (nodes at which no Dirichlet conditions are given).

Outline

Antil, H., Kouri, D. P., Lacasse, M.-D., and Ridzal, D., editors (2018).
Frontiers in PDE-constrained optimization, volume 163 of *The IMA Volumes in Mathematics and its Applications*.
Springer, New York.
Papers based on the workshop held at the Institute for Mathematics and its Applications, Minneapolis, MN, June 6–10, 2016.

Antil, H., Pfefferer, J., and Warma, M. (2017).
A note on semilinear fractional elliptic equation: analysis and discretization.
ESAIM Math. Model. Numer. Anal., 51(6):2049–2067.

Antil, H. and Warma, M. (2020).
Optimal control of fractional semilinear PDEs.
ESAIM Control Optim. Calc. Var., 26:Paper No. 5, 30.

Attouch, H., Buttazzo, G., and Michaille, G. (2006).
Variational analysis in Sobolev and BV spaces, volume 6 of *MPS/SIAM Series on Optimization*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
Applications to PDEs and optimization.

Braess, D. (1997).

Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics.
Cambridge University Press, Cambridge, New York.

Brenner, S. C. and Scott, L. (2008).
The Mathematical Theory of Finite Element Methods, volume 15 of *Texts in Applied Mathematics*.
Springer, New York, third edition.

Ciarlet, P. G. (2002).
The finite element method for elliptic problems, volume 40 of *Classics in Applied Mathematics*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

Ekeland, I. and Témam, R. (1999).
Convex analysis and variational problems, volume 28 of *Classics in Applied Mathematics*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition.
Translated from the French.

Ern, A. and Guermond, J. (2004).
Theory and practice of finite elements, volume 159 of *Applied Mathematical Sciences*.
Springer-Verlag, New York.

Heinkenschloss, M. (2008).
Numerical solution of implicitly constrained optimization problems.
Technical Report TR08–05, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892.
Available at `http://www.caam.rice.edu/~heinken/mh_publications.html`.

Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S. (2009).
Optimization with PDE constraints, volume 23 of *Mathematical Modelling: Theory and Applications*.
Springer, New York.

Ito, K. and Kunisch, K. (2008).
Lagrange multiplier approach to variational problems and applications, volume 15 of *Advances in Design and Control*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Lions, J.-L. (1971).
Optimal control of systems governed by partial differential equations.
Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York-Berlin.

Tröltzsch, F. (2010).

Optimal control of partial differential equations, volume 112 of *Graduate Studies in Mathematics*.
American Mathematical Society, Providence, RI.
Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

Ulbrich, M. (2011).
Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces, volume 11 of *MOS-SIAM Series on Optimization*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA.