



Lab course: Optimization with elliptic PDEs: Sheet 1

Exercise 1.1 (The linear elliptic test problem): We consider the linear quadratic elliptic test problem:

$$\min_{q \in L^2(\Omega), u \in H^1(\Omega)} J(q, u) = \frac{1}{2} \|u - u_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|q\|_{L^2(\Omega)}^2 \quad (1a)$$

$$\text{subject to } \begin{cases} -\Delta u + u = q & \text{in } \Omega, \\ \partial_n u = 0 & \text{on } \partial\Omega. \end{cases} \quad (1b)$$

The domain is the square $\Omega = (-1, 1) \times (-1, 1)$ and as desired states we consider either

$$u_d(x) = u_d^{\text{eig}}(x_1, x_2) = \sin\left(\frac{\pi}{2}x_1\right) \sin\left(\frac{\pi}{2}x_2\right), \quad (u_d^{\text{eig}})$$

$$\text{or } u_d^{\text{char}}(x) = \chi_{\{|x| \leq \frac{1}{2}\}}, \quad (u_d^{\text{char}})$$

where $|x| = \sqrt{x_1^2 + x_2^2}$ and χ is a characteristic function.

To solve the problem numerically, follow these steps:

1. Find the weak formulation of the PDE (1b) in the continuous and discrete setting (in terms of the variables $u \in H^1(\Omega)$ and $u_h \in V_h$). Formulate the discrete equation as a linear equation in terms of $\mathbf{u} \in \mathbb{R}^N$ (cf. sheet 0, Ex. 0.1).

2. Represent the discrete solution operator $S_h: \mathbf{q} \mapsto \mathbf{u}$ in terms of appropriate FE-matrices. Do the same for the discrete objective function $J_h: (\mathbf{q}, \mathbf{u}) \mapsto J_h(\mathbf{q}, \mathbf{u})$.

Hint: Use that $\|u_h - u_d\|^2 = \|u_h\|^2 - 2(u_h, u_d) + \|u_d\|^2$ and introduce the mass matrix.

3. In order to characterize the minimum, we have to write down the optimality conditions. To that purpose, we suggest first a “direct” method:

Start by formulating the reduced objective function

$$j_h(\mathbf{q}) = J_h(\mathbf{q}, S_h \mathbf{q})$$

in matrix-vector notation. Show that the derivatives of j_h can be represented as

$$j_h'(\mathbf{q})(\mathbf{x}) = \mathbf{x}'(\alpha \mathbf{M}_q + \mathbf{B}'(\mathbf{A}')^{-1}(\mathbf{M}_u \mathbf{A}^{-1} \mathbf{B} \mathbf{q} - \mathbf{f})), \quad (3)$$

$$j_h''(\mathbf{q})(\mathbf{x}, \mathbf{y}) = \mathbf{x}'(\alpha \mathbf{M}_q + \mathbf{B}'(\mathbf{A}')^{-1} \mathbf{M}_u \mathbf{A}^{-1} \mathbf{B}) \mathbf{y}, \quad (4)$$

for appropriate sparse matrices \mathbf{M}_q , \mathbf{M}_u , \mathbf{B} , \mathbf{A} , and a vector \mathbf{f} .

4. Argue that \mathbf{q} is optimal if (and only if) it is a solution of the linear equation

$$\left[\alpha \mathbf{M}_q + \mathbf{B}'(\mathbf{A}')^{-1} \mathbf{M}_u \mathbf{A}^{-1} \mathbf{B} \right] \mathbf{q} = \mathbf{B}'(\mathbf{A}')^{-1} \mathbf{f}. \quad (5)$$

Argue that the matrix on the left hand side is symmetric and positive definite.

5. We now want to solve the linear equation (5) with the cg-Method. Use the following steps:

- Assemble all required matrices and the vector \mathbf{f} .

Important: never directly compute the inverse \mathbf{A}^{-1} or $(\mathbf{A}')^{-1}$ or a full matrix-matrix product!

- Compute the right-hand side $\mathbf{g} = \mathbf{B}'(\mathbf{A}')^{-1}\mathbf{f}$.
- Implement an (anonymous) function that computes a matrix-vector product $\mathbf{H}\mathbf{x} = (\alpha\mathbf{M}_q + \mathbf{B}'(\mathbf{A}')^{-1}\mathbf{M}_u\mathbf{A}^{-1}\mathbf{B})\mathbf{x}$ for a given \mathbf{x} .
- Call the cg-Method (doc `pcg`) to solve $\mathbf{H}\mathbf{q} = \mathbf{g}$.

Hint: it is a good idea to use \mathbf{M}_q as a preconditioner for the pcg method. Why?

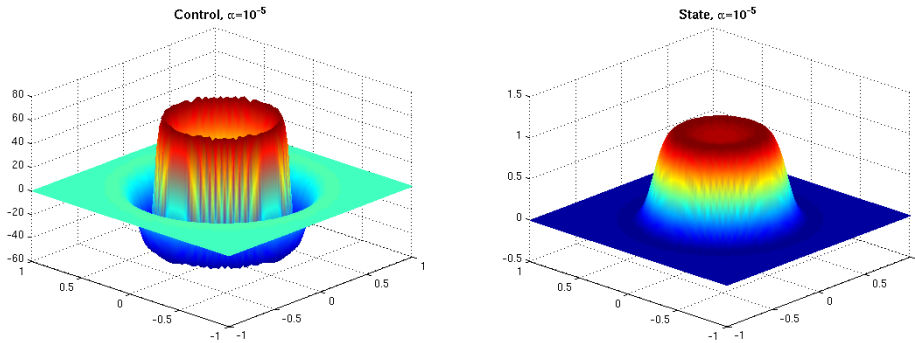


Figure 1: Optimal control and state for $\alpha = 10^{-5}$, desired state $u_d = u_d^{char}$.

6. For (u_d^{eig}) it is possible to compute the analytic solution of the problem (1a)–(1b) with arbitrary $\alpha > 0$. The optimal control is given by

$$q^{ref}(x_1, x_2) = \frac{\lambda}{1 + \alpha\lambda^2} \sin\left(\frac{\pi}{2}x_1\right) \sin\left(\frac{\pi}{2}x_2\right), \quad \text{where } \lambda = \left(1 + \frac{\pi^2}{2}\right).$$

Verify the correctness of the implementation by comparing the numerical with the analytic solution. Evaluate the L^2 error of the exact control on a sequence of successively refined meshes (cf. sheet 0, Ex. 0.1).

7. Compute the optimal solutions of the problem for (u_d^{char}) for regularization parameters $\alpha \in \{10^{-1}, 10^{-3}, 10^{-5}, 10^{-7}\}$. Explain why the discretization needs to be finer to accurately capture the behavior of the solutions for smaller α . Explain the behavior of the solutions for $\alpha \rightarrow 0$.

Bonus: Prove that the problem has no solution for (u_d^{char}) and $\alpha = 0$.

Bonus²: What is different in the case of (u_d^{eig}) ?

Exercise 1.2 (The test problem “reloaded”): We consider the same problem as in Ex. 1.1, but want to have another look at the discrete optimality system with a different approach: the Lagrangian formulation.

1. Formulate the Lagrange function to find the adjoint equation and optimality condition in the continuous and discrete setting. Write the system both in weak and in strong formulation in the former case and both in terms of (q_h, u_h, z_h) and in matrix-vector form with $(\mathbf{q}, \mathbf{u}, \mathbf{z})$ in the latter case.

Bonus: Show algebraically that the discrete optimality condition is equivalent to equation (5).

2. By writing all equations of the discrete optimality system in one big linear system, you should arrive at the “saddle point formulation”

$$\begin{pmatrix} \mathbf{M}_u & 0 & \mathbf{A}' \\ 0 & \alpha \mathbf{M}_q & \mathbf{B}' \\ \mathbf{A} & \mathbf{B} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{q} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ 0 \\ 0 \end{pmatrix}.$$

Assemble the correct matrices and load vector, build the block matrix and right-hand side and solve the resulting system (with the backslash operator).

Important: ensure that the block matrix is a sparse matrix (you can create a sparse zero matrix with `sparse`).

Exercise 1.3 (Convection-diffusion problem): Consider the optimization problem

$$\min_{q \in L^2(\Gamma_{in}), u \in H^1(\Omega)} J(q, u) = \frac{1}{2} \|u - u_d\|_{L^2(\Gamma_{out})}^2 + \frac{\alpha}{2} \|q\|_{L^2(\Gamma_{in})}^2 \quad (6a)$$

$$\text{subject to } \begin{cases} -\varepsilon \Delta u + b \cdot \nabla u = 0 & \text{in } \Omega, \\ \varepsilon \partial_n u = 0 & \text{on } \partial\Omega \setminus \Gamma_{in}, \\ \varepsilon \partial_n u + \rho u = \rho q & \text{on } \Gamma_{in}. \end{cases} \quad (6b)$$

We consider the unit square $\Omega = (0, 1) \times (0, 1)$. The boundary is divided into three parts, $\Gamma_{in} = \{0\} \times (0, 1)$ on the left, $\Gamma_{out} = \{1\} \times (0, 1)$ on the right, and the rest $\partial\Omega \setminus (\Gamma_{in} \cup \Gamma_{out}) = (0, 1) \times \{0, 1\}$. The velocity field b is assumed to be divergence free and the product $n \cdot b$ with the outer normal n is assumed to be positive on Γ_{out} , negative on Γ_{in} and zero on the rest. The coefficient ρ is given by $\rho = -n \cdot b$. Furthermore, $\varepsilon > 0$ is a small diffusion parameter.

In the first example the velocity field is explicitly given by $b = (1, 0)'$. Due to the special form of b we have $b \cdot \nabla u = \partial_{x_1} u$ in Ω and $\rho = -n \cdot b = 1$ on Γ_{in} .

In this example, u stands for a volume concentration. On the inflow boundary (where $n \cdot b < 0$) we add some volume in the form of the control q . This concentration is then transported to the right (according to b) and diffused depending on the magnitude of ε .

Our goal with this problem is to produce a desired concentration on the outflow by choosing an appropriate inflow concentration q . As desired state we consider

$$u_d^{\text{Gauss}}(x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_2 - \frac{1}{2})^2}{2\sigma^2}\right). \quad (u_d^{\text{Gauss}})$$

1. Solve the state equation for $q = u_d^{\text{Gauss}}$ for $\varepsilon \in \{1, 10^{-1}, 10^{-2}, 10^{-3}\}$ (cf. sheet 0, Ex 0.1). You can define multiple boundary pieces by modifying the definition of the mesh as follows:

```
boundary = Boundary('gamma_in', [...], 'gamma_out', [...], ..);
```

In the boundary related forms you have access to the characteristic function of a boundary piece evaluated at the quadrature points in the vector `data.chi.gamma_in` (etc.). This is useful since, e.g.,

$$(u, z)_{\Gamma_{in}} = \int_{\Gamma_{in}} uz \, ds = \int_{\partial\Omega} \chi_{\Gamma_{in}} uz \, ds.$$

2. Follow one of the methods from the previous exercises to solve the problem. In this example the space of the state and the control are not the same. The degrees of freedom on the boundary are only the nodes lying on Γ_{in} . You can find the corresponding indices of the nodes with the command `BoundaryNodes`. You can store the control in a vector that has all the nodal degrees of freedom of the mesh if you extend it by zero in the interior.

For the parameters you can try, e.g., $\sigma = 0.2$, $\alpha = 10^{-3}$ and $\varepsilon = 0.1$ as in Figure 2.

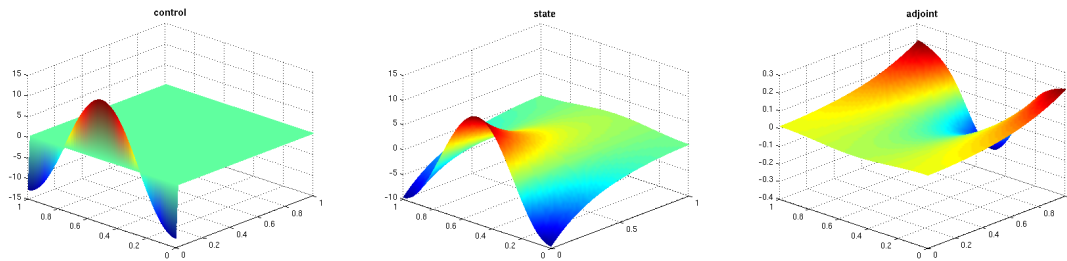


Figure 2: Optimal control, state and adjoint for $\sigma = 0.2$, $\alpha = 10^{-3}$, $\varepsilon = 0.1$.

3. Derive also the strong formulation of the adjoint equation. Give a physical interpretation of the adjoint equation and compare it with the state equation.

Hint: Use integration by parts.

4. *Bonus:* Use the strong formulation to identify a different approach of the discretization for the adjoint equation. Do they lead to the same result? Which formulation do you prefer?
5. *Bonus:* Modify your implementation for the parabolic velocity profile

$$b(x_1, x_2) = \begin{pmatrix} x_2(1 - x_2) \\ 0 \end{pmatrix}.$$

Exercise 1.4 (Bonus: An inverse problem): We consider again the problem (6), but on an L-shaped domain. The velocity profile is not given analytically here, but is created by solving the *Stokes equation* for a prescribed parabolic inflow. A plot of the geometric setup with the in- and outflow boundary Γ_{in} and Γ_{out} can be found in Figure 3.

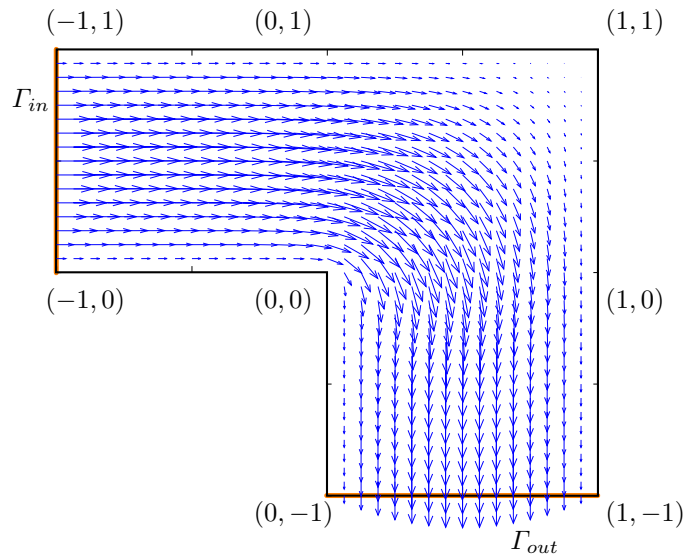


Figure 3: L-shaped domain and parabolic velocity profile

We choose the diffusion parameter as $\varepsilon = 0.002$. For an inflow q equal to (u_d^{Gauss}) with $\sigma = 0.2$ we get a simulation result as in Figure 4a.

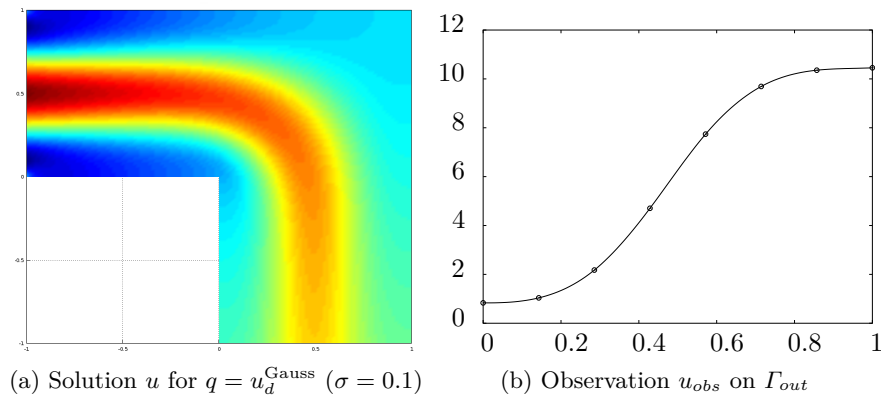


Figure 4: Simulation results.

However, from some other experiment we have recorded the observation u_{obs} on the outflow domain plotted in Figure 4b. It is the natural cubic spline interpolating the nodes:

x_1	0.00000	0.14286	0.28571	0.42857	0.57143	0.71429	0.85714	1.00000
$u_{obs}(x_1)$	0.83435	1.03768	2.17623	4.70531	7.73888	9.69081	10.35610	10.45425

- The domain and velocity profile can be produced with the command `[mesh, b1, b2] = SolveStokes(refine)`. Based on that, reproduce the simulation from Figure 4a. You have access to the outer normal on the boundary in the structure `data.normal`, i.e.,

`data.normal.x` for the x_1 component of the normal vector and `data.normal.y` for the x_2 component.

- You try to reconstruct the inflow q that has created Figure 4b. Therefore you use the problem formulation (6) from the previous exercise, which is now interpreted as a regularized least squares problem. The pointwise values of $u_d = u_{obs}$ can be evaluated with the commands `spline` and `ppval`.

Solve the problem for a decreasing sequence of regularization parameters α . Tabulate and plot the size of the *regularization term* $\frac{1}{2}\|q\|_{L^2(\Gamma_{in})}^2$ and the size of the *misfit term* $\frac{1}{2}\|u - u_d\|_{L^2(\Gamma_{out})}^2$. Plot the one against the other in logarithmic scale. What do you observe?

- Interpret the results: Which size of the regularization parameter appears optimal to you? Under the hypothesis that the inflow q was also a Gaussian, which mean value and variance do you suspect?