2.8. First-order necessary optimality conditions

Our starting point is the reduced problem

$$(*) \quad \min_{u \in U_{ad}} \ f(u) := \frac{1}{2} \| Su - y_d \|_H^2 + \frac{\lambda}{2} \| u \|_U^2 \ ,$$

our aim is to derive the optimality system and the KKT conditions, similar to the final-dimensional case.

Lemma 2.21. Let U be a real Banach space, $C \subset U$ a non-empty and convex subset. Let $f : U \to \mathbb{R}$ be Gâteaux-differentiable in an open subset of U containing C.

If $\bar{u} \in C$ is a solution of

$$\min_{u \in C} \ f(u)$$

then it holds the necessary <u>condition</u>

$$f'(\bar{u})(u - \bar{u}) \geq 0 \quad \text{for all } u \in C.$$

If f is convex, then this condition is also <u>sufficient</u>.

Proof: Let $u \in C$ be arbitrary. Since C convex,

$$\bar{u} + t(u - \bar{u}) \in C \ , \quad t \in (0,1].$$

Since $\bar{u}$ optimal, $f(\bar{u} + t(u - \bar{u})) \geq f(\bar{u})$, hence

$$\frac{1}{t} [f(\bar{u} + t(u - \bar{u})) - f(\bar{u})] \geq 0 \quad \text{for } t \in (0,1].$$

The limit $t \searrow 0$ yields $f'(\bar{u})(u - \bar{u}) \geq 0$.

Now suppose $\bar{u}$ solves the variational inequality and f is convex, hence for $t \in [0,1]$

$$f(\underbrace{tu + (1-t)\bar{u}}_{= \bar{u} + t(u-\bar{u})}) \leq t f(u) + (1-t) f(\bar{u})$$

$$\Longleftrightarrow f(u) - f(\bar{u}) \geq \frac{1}{t}\left(f(\bar{u} + t(u-\bar{u})) - f(\bar{u}) \right)$$

In the limit $t \searrow 0$, it holds

$$f(u) - f(\bar{u}) \geq f'(\bar{u})(u-\bar{u}) \geq 0 \quad \text{for all } u \in C$$

since the variational inequality holds.

Therefore, $f(u) \geq f(\bar{u})$ for all $u \in C$, which proves the claim. ⊠

Remark: The result remains true, if merely the existence of all directional derivatives of f is assumed.

Theorem 2.22: Let U, H be real Hilbert spaces, $y_d \in H$ given, $U_{ad} \subset U$ non-empty and convex, and $\lambda \geq 0$. Moreover, let $S: U \to H$ denote a continuous linear operator.

Then $\bar{u} \in U_{ad}$ is a solution to (∗) if and only if $\bar{u}$ solves the variational inequality

$$\left(S^*(S\bar{u} - y_d) + \lambda \bar{u}, \, u - \bar{u} \right)_U \geq 0 \quad \text{for all } u \in U_{ad}.$$

Proof: From the examples on p.43 we know that the functional $f(u)$ in (∗) is Fréchet-differentiable with $f'(\bar{u}) = S^*(S\bar{u} - y_d) + \lambda \bar{u}$. The claim then follows directly using Lemma 2.21. ⊠

We can apply these results to various optimal control problems.

Optimal stationary heat source

$$(**) \begin{cases} \min \quad J(y,u) := \frac{1}{2} \|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2} \|u\|^2_{L^2(\Omega)} \\[2mm] \text{s.t.} \quad (***) \quad -\Delta y = \beta u \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma \\[2mm] \text{and} \quad u_a(x) \leq u(x) \leq u_b(x) \text{ for a.e. } x \in \Omega. \end{cases}$$

As above, we denote by $S: L^2(\Omega) \to L^2(\Omega)$ the solution operator of the boundary value problem. Using Theorem 2.22, any optimal control $\bar{u}$ must obey the variational inequality

$$\left(S^*(S\bar{u} - y_\Omega) + \lambda \bar{u}, u - \bar{u} \right)_{L^2(\Omega)} \geq 0 \text{ for all } u \in U_{ad}.$$

Herein, S^* is the adjoint operator to S, which is yet to be determined. For this, we require the following auxilliary result.

Lemma 2.23: Let $u, z \in L^2(\Omega)$ and $c_0, \beta \in L^\infty(\Omega)$ with $c_0 \geq 0$ a.e. in Ω be given functions. Let y and p denote weak solutions of

$$\begin{cases} -\Delta y + c_0 y = \beta u \text{ in } \Omega, \\ \qquad\qquad y = 0 \quad \text{on } \Gamma, \end{cases}$$

and

$$\begin{cases} -\Delta p + c_0 p = z \quad \text{in } \Omega, \\ \qquad\qquad p = 0 \quad \text{on } \Gamma, \end{cases}$$

respectively. Then $\displaystyle\int_\Omega z y \, dx = \int_\Omega \beta p u \, dx.$

Proof: Consider the weak formulations with $p \in H_0^1(\Omega)$ and $y \in H_0^1(\Omega)$ as test functions, respectively:

$$\int_\Omega \nabla y \cdot \nabla p + c_0 y p \, dx = \int_\Omega \beta u p \, dx \quad \text{and} \quad \int_\Omega \nabla p \cdot \nabla y + c_0 p y \, dx = \int_\Omega z y \, dx.$$

Therefore, the right hand sides must be identical. ∎

<u>Lemma 2.24</u>: For the boundary value problem (***), the adjoint operator $S^*: L^2(\Omega) \longrightarrow L^2(\Omega)$ is given by

$$S^* z = \beta p,$$

where $p \in H_0^1(\Omega)$ is the weak solution of

$$\begin{cases} -\Delta p = z & \text{in } \Omega, \\ \quad\; p = 0 & \text{on } \Gamma. \end{cases}$$

<u>Proof</u>: By the definition of the adjoint operator, the identity $(z, Su)_{L^2(\Omega)} = (S^* z, u)_{L^2(\Omega)}$ for all $u, z \in L^2(\Omega)$ must hold. Using Lemma 2.23 with $c_0 \equiv 0$ and $y = Su$, we find: $(z, Su)_{L^2(\Omega)} = (z, y)_{L^2(\Omega)} = (\beta p, u)_{L^2(\Omega)}$.

By our well-posedness results for elliptic boundary value problems, the mapping $z \longmapsto \beta p$ is a continuous linear operator from $L^2(\Omega)$ into $L^2(\Omega)$.

Since z, u are arbitrary, and S^* is uniquely defined, we conclude that $S^* z = \beta p$. $\boxtimes$

<u>Remark</u>: (i) This approach to obtain the adjoint operator is not very intuitive, yet well-defined. We will later introduce a "formal Lagrange method", which is an effective tool to find the adjoint operator S^* and the specific PDE for it.

(ii) It may now be clearer why we worked with $S: L^2(\Omega) \longrightarrow L^2(\Omega)$ instead of $G: L^2(\Omega) \longrightarrow H_0^1(\Omega)$. If we work with G, we would now encounter $G^*: (H_0^1(\Omega))^* \longrightarrow L^2(\Omega)$, which would involve the dual space $(H_0^1(\Omega))^*$. This is possible, and allows a bit more general problems (e.g. right hand sides in $(H_0^1(\Omega))^*$), but at the same time may be more complicated than to work in $L^2(\Omega)$, cf. Section 2.13 in Tröltzsch's book.

The variational inequality can easily be rewritten if S^* is known.

Definition : The weak solution $p \in H_0^1(\Omega)$ to the
adjoint equation

$$\begin{cases} -\Delta p = \bar{y} - y_\Omega & \text{in } \Omega, \\ p = 0 & \text{on } \Gamma, \end{cases}$$

is called adjoint state associated with $\bar{y}$.

Note that $(y - y_\Omega) \in L^2(\Omega)$, hence the adjoint equation admits a unique solution $p \in H_0^1(\Omega)$.

Setting now $z = \bar{y} - y_\Omega$, we employ Lemma 2.24 to obtain $S^*(S\bar{u} - y_\Omega) = S^*(\bar{y} - y_\Omega) = \beta p$ (p solution of the adjoint equation), and hence

$$(\beta p + \lambda \bar{u}, u - \bar{u})_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{ad}.$$

Hence, a control u is optimal, together with associated state y and adjoint state p, for the problem (**) if and only if (that it is also sufficient follows from the convexity of the reduced functional $f(\cdot)$) the following optimality system is satisfied:

$$(OS) \quad \begin{cases} -\Delta y = \beta u & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma, \end{cases} \quad , \quad u \in U_{ad}$$

$$\begin{cases} -\Delta p = y - y_\Omega & \text{in } \Omega \\ p = 0 & \text{on } \Gamma \end{cases}$$

$$(\beta p + \lambda u, v - u)_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{ad}$$