R. Herzog, G. Müller
Universität Heidelberg

Exercise 8 (Solution)

Date issued: 2nd December 2024

Homework Problem 8.1. (Compact operators in the direct method of variational calculus)

Let X, Y and Z be Banach spaces. A linear operator $A\colon X \to Y$ is called **compact** if it maps bounded sets to sets whose closure is compact.

(a) Show that $A\colon X \to Y$ is compact if and only if the sequence of images $A(x^{(k)})$ in Y for any a bounded sequence $x^{(k)}$ in X has a convergent subsequence.

(b) Show that if $A\colon X \to Y$ is compact, then A is continuous.

(c) Show that if A is compact, then for any $B \in \mathcal{L}(Y, Z)$ the operators $B \circ A$ and $A \circ B$ are compact.

(d) Explain how compactness of an operator can play a role in the proof of the existence of optimizers for optimization problems of the type (5.8) when applying the direct method of variational calculus.

Solution.

(a) Let A be linear and compact as defined. Further, let $x^{(k)}$ be a bounded sequence in X. Then the set of images $I \coloneqq \{A(x^{(k)} \mid k \in \mathbb{N}\}$ has compact closure, i. e., any sequence in the close of the set has a convergent subsequence with a limit point in $\overline{I}$, this especially applies to the sequence of images itself.

Now for the converse, let A map bounded sequences to those with convergent subsequences. Let U be a bounded subset of X and consider a sequence $y^{(k)}$ in $\overline{A(U)}$. If there is a subsequence with preimage sequence such that $y^{(k^{(l)})} = A\left(x^{(k^{(l)})}\right)$ for $x^{(k^{(l)})}$ in U, which is also bounded of course, we immediately obtain that $A\left(x^{(k^{(l)})}\right)$ has a convergent subsequence. Otherwise, due to the definition of the closure, we can choose a sequence $\widetilde{y^{(k)}}$ with the required preimage property and $\|\widetilde{y^{(k)}} - y^{(k)}\|_Y \leqslant \frac{1}{k}$ which provides the convergent subsequence in $\overline{A(U)}$ and therefore compactness of $\overline{A(U)}$.

(b) We show boundedness of A. We assume that A is not the zero map, otherwise, there is nothing to show. Now, consider a sequence $x^{(k)}$ with

$$\frac{\|Ax^{(k)}\|_Y}{\|x^{(k)}\|_X} \to \sup_{x \in X \setminus \{0\}} \frac{\|Ax\|_Y}{\|x\|_X}.$$

Then the set $\{\frac{x^{(k)}}{\|x^{(k)}\|} \mid k \in \mathbb{N}\}$ is bounded and hence their images contain a convergent subsequence, meaning that the sup is finite and coincides with the norm of A.

(c) If A is compact and B is continuous, then B is bounded and maps bounded sets to bounded sets, which A maps to precompact sets, i. e., $B \circ A$ is compact. Additionally, since A maps bounded sets to precompact sets, $A(U)$ for bounded $U \subseteq X$ is precompact and since B is continuous, $B \circ A$ is compact as well.

(d) In the DMVC, we have shifted some of the work of showing nice properties of the admissible set (compactness of the set, in the best case) to having to show stronger results for the functional, i. e., weak sequential l.s.c. We know that convexity and strong continuity yield w.s.l.s.c., but for nonconvex functionals, we are not well prepared. Often times, one can identify a target functional structure of the type $f(u) = g(A(u))$, for a continuous functional g mapping some space Y to $\mathbb{R}$ and $A \colon X \to Y$ a compact operator. If A is compact (e.g., the embedding of $W^{1,2}$ into L^2 on bounded Lipschitz domains) it is sufficient for g to be strongly l.s.c. in order to obtain the desired result.

Homework Problem 8.2. (Optimizer invariance for control-reduced problems)

Suppose that Y and U are normed linear spaces. Show Lemma 6.1, i. e., the following statements:

(a) Suppose that $G \colon U_{\text{ad}} \to Y$ provides, for any $u \in U_{\text{ad}}$, the unique solution $y = G(u)$ of the constraint $e(y, u) = 0$.

  (i) If (y^*, u^*) is a global minimizer of the original (6.1), then u^* is a global minimizer of the reduced problem (6.2).

  (ii) If u^* is a global minimizer of the reduced problem (6.2), then $(G(u^*), u^*)$ is a global minimizer of the original problem (6.1).

(b) Suppose in addition that $G \colon U_{\text{ad}} \to Y$ is continuous on U_{ad}.

  (i) If (y^*, u^*) is a local minimizer of the original (6.1), then u^* is a local minimizer of the reduced problem (6.2).

(ii) If u^* is a local minimizer of the reduced problem (6.2), then $(G(u^*), u^*)$ is a local minimizer of the original problem (6.1).

Solution.

(a) We prove this on indirectly.

(i) Suppose there were a $\bar{u} \in U_{\text{ad}}$ such that $f(\bar{u}) < f(u^*)$ in the reduced problem (6.2). Then we can apply G to $\bar{u}$ and obtain the state $\bar{y} \coloneqq G(\bar{u})$ and

$$J(\bar{y}, \bar{u}) = J(G(\bar{u}), \bar{u}) = f(\bar{u}) < f(u^*) = J(G(u^*), u^*) = J(y^*, u^*).$$

Since the pair $(\bar{y}, \bar{u})$ is feasible by assumption on G, this shows suboptimality of (y^*, u^*) in the original problem and therefore a contradiction.

(ii) Suppose there were a feasible pair $(\bar{y}, \bar{u})$ with $J(\bar{y}, \bar{u}) < J(y^*, u^*)$ in the original problem (6.1). Then by definition of feasibility, $\bar{u} \in U_{\text{ad}}$ and $G(\bar{u}) = \bar{y}$. Accordingly, feasiblity of (y^*, u^*) shows that

$$f(\bar{u}) = J(G(\bar{u}), \bar{u}) = J(\bar{y}, \bar{u}) < J(y^*, u^*) = J(G(u^*), u^*) = f(u^*),$$

i. e., suboptimality of u^* in the reduced problem and therefore a contradiction.

(b) We prove this one directly and assume that $\|\cdot\|_{Y\times U} = \|Y\| + \|U\|$.

(i) Let (y^*, u^*) be locally optimal for the original problem. Then there is an $\varepsilon > 0$ such that $J(y^*, u^*) \leqslant J(y, u)$ for all feasible $(y, u) \in B_\varepsilon((y^*, u^*))$. Now there is a $\delta > 0$ with $\delta < \frac{1}{2}\varepsilon$ such that $\|G(u) - G(u^*)\|_Y \leqslant \frac{1}{2}\varepsilon$ for all feasible $u \in B_\delta(u^*)$ and therefore $(G(u), u) \subseteq B_\varepsilon((y^*, u^*))$ for all feasible $u \in B_\delta(u^*)$, which shows local optimality of u^* in the reduced problem because $(G(u), u)$ is feasible for $u \in U_{\text{ad}}$ by definition of G.

(ii) Let $u^* \in U_{\text{ad}}$ be locally optimal for the reduced problem. Then there is an $\varepsilon > 0$ such that $f(u^*) \leqslant f(u)$ for all feasible $u \in B_\varepsilon(u^*)$. Now let $(y, u) \in B_\varepsilon((G(u^*), u^*))$, be a feasible point, then of course $u \in U_{\text{ad}} \cap B_\varepsilon(u^*)$ and therefore

$$J(y, u) = J(G(u), u) = f(u) \geqslant f(u^*) = J(G(u^*), u^*) = J(y^*, u^*).$$

You are not expected to turn in your solutions.