Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4183 Opt. II
Spring 2017

Exercise set 1

Reading material: Chapter 1 & Section 2.1–2.2 from [Tröltsch].

1 We consider a(n artificial) finite-dimensional optimal control problem for $y \in \mathbb{R}^2$ with a control parameter $u \in \mathbb{R}$.

The state equation is:

$$\begin{aligned} y_1 + y_2 &= u, \\ y_2 &= 2u, \end{aligned} \tag{1}$$

and the const functional is

$$J(y, u) = \frac{1}{2}[(y_1 - 1)^2 + (y_2 - 2)^2] + \frac{\lambda}{2}u^2, \tag{2}$$

where $\lambda > 0$.

a) Derive the explicit expressions for the reduced cost functional and its gradient.

Solution: The control-to-state operator $y = Su$ is obtained by solving the state equations yeilding $S = [-1, 2]^T$. The reduced cost function and its gradient are:

$$\begin{aligned} f(u) &= J(Su, u) = \frac{5+\lambda}{2}u^2 - 3u + \frac{5}{2}, \\ f'(u) &= (5+\lambda)u - 3. \end{aligned}$$

b) Formulate the adjoint problem and compute the reduced gradient with the help of the adjoint state.

Solution: The state equation in the matrix-vector form can be stated as

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{=:B} u.$$

He adjoint system is then $A^T p = \nabla_y J$, or

$$\begin{aligned} p_1 &= y_1 - 1, \\ p_1 + p_2 &= y_2 - 2, \end{aligned}$$

thus $p_2 = -y_1 + y_2 - 1$. Finally, the reduced gradient is

$$\begin{aligned} f'(u) &= B^T p + \nabla_u J = 1(y_1 - 1) + 2(-y_1 + y_2 - 1) + \lambda u \\ &= -u - 1 + 2(u + 2u - 1) + \lambda u = (5+\lambda)u - 3. \end{aligned}$$

c) Assuming $U_{\mathrm{ad}} = \mathbb{R}$ state the first order necessary optimality conditions for this problem.

Solution: In the absense of restrictions on the control the first order necessary optimality conditions are

$$\begin{aligned} Ay &= Bu \\ A^T p &= \nabla_y J \\ \underbrace{B^T p + \nabla_u J}_{=f'(u)} &= 0. \end{aligned}$$

These can even be solved, namely $u = 3/(5+\lambda)$ etc.

2 Let V be a normed space and V^* be its dual. Verify the fact that the expression

$$\|f\|_{V^*} = \sup_{x \in V \setminus \{0\}} \frac{|f(x)|}{\|x\|_V}$$

defines a norm on V^*.

Solution: (a): separation. If $\|f\|_{V^*} = 0$ then $\forall x \in V \setminus \{0\}$: $|f(x)| = 0$. Since f is linear, it means that $f(0) = f(x-x) = f(x) - f(x) = 0$. Therefore $\forall x \in V : f(x) = 0$, or $f = 0$ in V'.

(b): absolute homogeneity. For any $f \in V'$, any $\alpha \in \mathbb{R}$, and any $x \in V$ we have: $|(\alpha f)(x)| = |\alpha f(x)| = |\alpha||f(x)|$. As a result, $\|\alpha f\|_{V'} = |\alpha|\|f\|_{V'}$.

(c): triangle inequality. For any $x \in V \setminus \{0\}$, and any $f, g \in V'$ we can write $|(f+g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)|$. Therefore we can take the supremum on both sides of this inequality:

$$\sup_{x \in V \setminus \{0\}} \frac{|(f+g)(x)|}{\|x\|_V} \leq \sup_{x \in V \setminus \{0\}} \left[\frac{|f(x)|}{\|x\|_V} + \frac{|g(x)|}{\|x\|_V} \right]$$

Finally

$$\begin{aligned} \sup_{x \in V \setminus \{0\}} \left[\frac{|f(x)|}{\|x\|_V} + \frac{|g(x)|}{\|x\|_V} \right] &= \sup_{x = y \in V \setminus \{0\}} \left[\frac{|f(x)|}{\|x\|_V} + \frac{|g(y)|}{\|y\|_V} \right] \leq \sup_{x, y \in V \setminus \{0\}} \left[\frac{|f(x)|}{\|x\|_V} + \frac{|g(y)|}{\|y\|_V} \right] \\ &= \sup_{x \in V \setminus \{0\}} \frac{|f(x)|}{\|x\|_V} + \sup_{y \in V \setminus \{0\}} \frac{|g(y)|}{\|y\|_V} \end{aligned}$$

where the inequality holds because we take supremum over a larger set on the right (we drop the requirement $x = y$).

3 Let H be a Hilbert space, and consider an arbitrary $y \in H$. Show that the function $f(x) = (x, y)$ defines a bounded linear functional.

Solution: f is linear because the inner product is bilinear. f is bounded owing to Cauchy-Schwarz inequality: $|f(x)| = |(x, y)| \leq \|x\|_H \|y\|_H$. As a result, $\|f\|_{H'} \leq \|y\|_H$.[1][2]

[1](In fact $|f(y)| = \|y\|_H^2$. Therefore from the definition of $\|f\|_H'$ one can easily see that $\|f\|_{H'} = \|y\|_H$.

[2]The map $R : H \to H'$ given by $y \mapsto f(\cdot) = (x, \cdot)$ is called the Riesz map.

4 Let H be a Hilbert space, and consider an arbitrary $f \in H^* \setminus \{0\}$. Let us define $C = \{\, x \in H \mid f(x) = 1 \,\}$.

a) Show that C is a non-empty closed convex set.

Solution: (i) Since $f \neq 0$ it follows that there is $\hat{x} \in H$: $f(\hat{x}) \neq 0$. Then $f(\hat{x}/f(\hat{x})) = f(\hat{x})/f(\hat{x}) = 1$, and therefore $\hat{y} = \hat{x}/f(\hat{x}) \in C$.
(ii) Let $\{x_k\}_{k=1}^{\infty} \in C$ be a Cauchy sequence. Since H is complete, the sequence has a limit $\hat{x} \in H$. $|1-f(\hat{x})| = |f(x_k)-f(\hat{x})| = |f(x_k-\hat{x})| \leq \|f\|_{H'}\|x_k-\hat{x}\|_H \to 0$ as $k \to \infty$. As a consequence $f(\hat{x}) = 1$, $\hat{x} \in C$, and C is closed.
(iii) Let $x, y \in C$ and $0 \leq \lambda \leq 1$. $f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y) = \lambda + (1-\lambda) = 1$, and $\lambda x + (1-\lambda)y \in C$. Thus C is convex.

b) Let $\hat{y} \in H$ be an arbitrary vector in C. Show that $C = \hat{y} + \ker f$, where $\ker f = \{\, x \in H \mid f(x) = 0 \,\}$.

Solution: Take an arbitrary $x \in C$. Then $f(x-\hat{y}) = f(x) - f(\hat{y}) = 1 - 1 = 0$ and $x - \hat{y} \in \ker f$. Therefore $C - \hat{y} \subset \ker f$.
Similarly, take any $z \in \ker f$. Then $f(\hat{y} + z) = f(\hat{y}) + f(z) = 1 + 0 = 1$. Therefore $\hat{y} + \ker f \subset C$.

c) Let $\bar{y} \in H$ be the shortest vector in C, that is, $\bar{y} = \arg\min_{y \in C} \|y\|_H^2$. Show that $\bar{y} \perp \ker f$, that is, $(\bar{y}, z) = 0$ for all $z \in \ker f$. Hint: consider perturbations of $\bar{y} \pm \delta z$, where $\delta \in \mathbb{R}$ and $z \in \ker f$. Use the optimality of $\bar{y}$.

Solution: Take any $z \in \ker f$ and $\delta > 0$. Then $\pm\delta z \in \ker f$ as well. Owing to the previous point (superposition principle) $\bar{y} \pm \delta z \in C$, and therefore $\|\bar{y} \pm \delta z\|_H^2 = (\bar{y} \pm \delta z, \bar{y} \pm \delta z) = \|\bar{y}\|_H^2 \pm 2\delta(\bar{y}, z) + \delta^2\|z\|_H^2 \geq \|\bar{y}\|_H^2$, as $\bar{y}$ is the shortest vector in C. We end up with the inequality

$$\pm 2(\bar{y}, z) + \delta\|z\|_H^2 \geq 0.$$

By letting $\delta \to 0$ we can see that the only possibility for this inequality to hold is for $(\bar{y}, z) = 0$.

d) Show that $f(x) = (\tilde{y}, x)$, where $\tilde{y} = \bar{y}/\|\bar{y}\|^2$. Hint: consider two cases: $x \in \ker f$ and $x \notin \ker f$. In the latter case $x/f(x) \in C$, to which the result from **b)** can be applied.

Solution: First of all note that $\bar{y} \neq 0$ because $f(\bar{y}) = 1$ and f is linear.
Let $\tilde{y}$ be as above. If $x \in \ker f$ then $(\bar{y}, x) = 0$ by **c)** and therefore $0 = f(x) = (\bar{y}, x)/\|\bar{y}\|_H^2$.
If $x \notin \ker f$ then $x/f(x) \in C$, owing to the linearity of f. Therefore $x/f(x) = \bar{y} + z$, where $z \in \ker f$. As a consequence we have $(\tilde{y}, x) = (\bar{y}/\|\bar{y}\|_H^2, f(x)\bar{y} + f(x)z) = f(x)(\bar{y}, \bar{y})/\|\bar{y}\|_H^2 + f(x)(\bar{y}, z)/\|\bar{y}\|_H^2 = f(x) \cdot 1 + f(x) \cdot 0 = f(x)$.

e) Show that $\|f\|_{H^*} = \|\tilde{y}\|_H$.

Solution:

$$\|\tilde{y}\|_H = \frac{|(\tilde{y}, \tilde{y})|}{\|\tilde{y}\|_H} \leq \sup_{x \in H\setminus\{0\}} \frac{|(\tilde{y}, x)|}{\|x\|_H} \leq \sup_{x \in H\setminus\{0\}} \frac{\|\tilde{y}\|_H \|x\|_H}{\|x\|_H} = \|\tilde{y}\|_H,$$

where the second inequality is owing to Cauchy–Schwarz.

The last exercise is known as the Riesz representation theorem, which constructs an isometry from H^* into H.