Exercise Sheet 10.

Hand in due to: **Friday, 02/06/2023, 14:00**

Exercise 1 (Compact operators | 4 points).

Let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be an operator defined by $Tx = (\lambda_n x_n)_{n\geq 1}$ for a sequence $\lambda_n \to 0$. Prove that T is compact.

Hint. *To show this, approximate T by operators T_N such that $T_N x = (\lambda_1 x_1, \dots, \lambda_N x_N, 0, \dots)$. Prove then that T_N are compact operators and that the limit operator of compact operators is also compact.*

Exercise 2 (Space of Hilbert-Schmidt operators | 4 points).

Let A and B be two Hilbert-Schmidt operators that act on a Hilbert space H. We define the Hilbert-Schmidt inner product and the related norm as

$$\langle A, B\rangle_{HS} := \sum_{i\in I} \langle Ae_i, Be_i\rangle_H \quad \text{and} \quad \|A\|_{HS}^2 = \sum_{i\in I} \|Ae_i\|_H^2,$$

where $\{e_i : i \in I\}$ is an orthonormal basis and I is an index set. Show the following properties:

(a) The Hilbert-Schmidt norm does not depend on the choice of the orthonormal basis.

(b) If $T : H \to H$ is a bounded linear operator, then we have $\|T\| \leq \|T\|_{HS}$.

(c) Consider the integral operator $T_K u = \int_\Omega K(x, y)u(y)dy$ for $K \in L^2(\Omega\times\Omega)$. Then, there holds $\|T_K\|_{HS} = \|K\|_{L^2(\Omega\times\Omega)}$.

(d) The composition of two Hilbert-Schmidt operators is again a Hilbert-Schmidt operator.

Exercise 3 (Canonical separating hyperplane | 4 points).

Let a dataset $D := (X_N, \{y_i\}_{i=1}^N) \subset \Omega \times \{1, 1\}^N$, where $\Omega \subset \mathbb{R}^d$, be linearly separable in $\mathbb{R}^d$, i.e., there exist $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$ such that $y_i(\langle w, x_i\rangle + b) > 0,\ 1 \leq i \leq N$. We then define a separating hyperplane $H(w, b) := \{x \in \mathbb{R}^d : \langle w, x\rangle + b = 0\}$. Next, we formulate classification problem as

$$\min_{w\in\mathbb{R}^d, b\in\mathbb{R}} \frac{1}{2}\|w\|_2^2 \text{ such that } y_i(\langle w, x_i\rangle + b) > 0,\ 1 \leq i \leq N. \tag{1}$$

Prove the following statements:

(a) The margin can be computed as

$$\gamma := \operatorname{dist}\big(X_N, H(w, b)\big) = \min_{1\leq i\leq N} \frac{y_i(\langle w, x_i\rangle + b)}{\|w\|_2^2}$$

and $\gamma > 0$.

(b) We call a separating hyperplane $H(w, b)$ canonical, if w and b are scaled such that $\gamma = 1/\|w\|^2$. Let (w^*, b^*) be a solution of classification problem (1). Then, $H(w^*, b^*)$ is a canonical separating hyperplane.

Exercise 4 (Dual problem | 4 points).

Let a dataset $D := (X_N, \{y_i\}_{i=1}^N) \subset \Omega \times \{1,1\}^N$, where $\Omega \subset \mathbb{R}^d$, be linearly separable in $\mathbb{R}^d$. We formulate dual classification problem in accordance with

$$\max_{\alpha \in \mathbb{R}^N} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \text{ such that } \sum_{i=1}^N \alpha_i y_i = 0 \text{ and } \alpha_i \geq 0 \text{ for all } i. \quad (2)$$

Let α^* be a solution of (2). Define $w^* \in \mathbb{R}^d$ by $w^* := \sum_{i=1}^N \alpha_i^* y_i x_i$ and $b^* \in \mathbb{R}$ such that, for an arbitrary i with $\alpha_i^* \neq 0$, there holds $y_i(\langle w, x_i \rangle + b) = 1$. Show that (w^*, b^*) is a solution of primal problem (1).

Hint. *Prove that* (w^*, b^*) *satisfy the Karush-Kuhn-Tucker conditions.*