3 Compact Operators, Generalized Inverse, Best-Approximate Solution

As we have already heard in the lecture a mathematical problem is well - posed in the sense of Hadamard if the following properties hold true

- For all admissible data, a solution exists
- For all admissible data, the solution is unique
- The solution depends continuously on the data

If one of these properties is not fulfilled, we speak of an ill-posed problem. In the sequel we consider linear operator equations of the form

$$Kx = y, \quad K : X \to Y, \tag{22}$$

where X and Y are Hilbert spaces. Since a variety of inverse problems (see 1D backwards heat equation) results in the formulation of an integral equation of the first kind

$$\int_G k(s,t)x(t)dt = y(s) \tag{23}$$

we have this kind of problem in the following in mind. Here $x(t)$ is the searched quantity, $k(s,t) \in L^2(G \times G)$ denotes the kernel and $y \in L^2(G)$ the input, e.g. some measured data.

Since integral operators are (under suitable conditions) "compact operators" we first review some basic facts about compact operators. Remember that the linear operator $K : X \to$ between Hilbert spaces X and Y is called compact if for all bounded sets $B \subseteq X, \overline{K(B)}$ is compact.

For any compact operator K between Hilbert spaces X and Y, there exists a singular system $(\sigma_i, u_i, v_i)_{i \in \mathbb{N}}$ which is defined as follows. If K^* denotes the adjoint of K (recall $\forall x \in X$ and $\forall y \in Y$ $\langle Kx, y \rangle = \langle x, K^* y \rangle$ holds), then $(\sigma_i)_{i \in \mathbb{N}}$ are the non-zero eigenvalues of the self-adjoint operator K^*K (and also of KK^*) and are ordered in decreasing order. The $(u_i)_{i \in \mathbb{N}}$ are a complete orthonormal system of eigenvalues of K^*K spanning $\overline{R(K^*)} = \overline{R(K^*K)}$ and the $(v_i)_{i \in \mathbb{N}}$ are defined via

$$v_i := \frac{Ku_i}{||Ku_i||}.$$

The $(v_i)_{i \in \mathbb{N}}$ are a complete orthonormal system of eigenvalues of KK^* and span $\overline{R(K)} = \overline{R(KK^*)}$. The following formulas hold

$$Ku_i = \sigma v_i \tag{24}$$

$$K^* v_i = \sigma_i u_i \tag{25}$$

$$Kx = \sum_{i=1}^{\infty} \sigma_i \langle x, u_i \rangle v_i \quad (x \in X) \tag{26}$$

$$K^* y = \sum_{i=1}^{\infty} \sigma_i \langle y, v_i \rangle u_i \quad (y \in Y). \tag{27}$$

The latter two sums are called "singular value expansion" (SVE) and converge in the Hilbert space norms of X and Y, respectively. If there are infinitely many singular values they accumulate (only) at zero, i.e.

$$\lim_{i\to\infty} \sigma_i = 0$$

which is a crucial fact for the ill-posedness of integral - equations of the first kind.
We will now relax the notion of a solution. In general for any operator equation $Tx = y$ we have existence (1) of a solution if $y \in R(T)$ holds and uniqueness (2) if $\mathcal{N}(T) = \{0\}$. Stability is obtained, assuming that (1) and (2) hold, so that T^{-1} exists, if T^{-1} is bounded. Always assuming that (1) and (2) are fulfilled would be very restrictive, therefore, we are interested in some generalized notion of a solution, that approximates x in $Tx = y$ in a unique way.

Definition 1 *Let $T : X \to Y$ be a bounded linear operator.*

a) $x \in X$ is called "least-squares solution" of $Tx = y$ if

$$||Tx - y|| = \inf_{z\in X} \{||Tz - y|| \mid z \in X\}$$

b) $x \in X$ is called "best approximate solution" of $Tx = y$ if x is least-squares solution and

$$||x|| = \inf_{z\in X} \{||z|| \mid z \text{ is least-squares solution of } Tx = y\}\,.$$

The best-approximate solution is therefore the least-squares solution with minimal norm and is closely related to the "Moore-Penrose" (generalized) inverse $T^\dagger$ of T, which will turn out to be the solution operator mapping y onto the best-approximate solution of $Tx = y$. We define the Moore-Penrose inverse in an operator theoretic way by restricting the domain and range of T in such a way that the resulting restricted operator is invertible, its inverse will then be extended to its maximal domain.

Definition 2 *The Moore-Penrose inverse $T^\dagger$ of $T \in \mathcal{L}(X, Y)$ is defined as the unique linear extension of $\tilde{T}^{-1}$ to*

$$\mathcal{D}(T^\dagger) := \mathcal{R}(T) + \mathcal{R}(T)^\perp \tag{28}$$

with

$$\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp \tag{29}$$

where

$$\tilde{T} := T_{|\mathcal{N}(T)^\perp} : \mathcal{N}(T)^\perp \to \mathcal{R}(T). \tag{30}$$

$T^\dagger$ is well defined: Since $\mathcal{N}(\tilde{T}) = \{0\}$ and $\mathcal{R}(\tilde{T}) = \mathcal{R}(T)$ the inverse $\tilde{T}^1$ exists. Due to (29) and the requirement that $T^\dagger$ is linear for any $y \in \mathcal{D}(T^\dagger)$ with the unique representation $y = y_1 + y_2$, where $y_1 \in \mathcal{R}(T)$ and $y_2 \in \mathcal{R}(T)^\perp$, $T^\dagger y$ has to be $\tilde{T}^{-1} y_1$.

Proposition 1 *Let now P and Q be orthogonal projectors onto $\mathcal{N}(T)$ and $\mathcal{R}(T)$, respectively. Then,* *a) the "Moore-Penrose equations"*

$$\begin{aligned} TT^\dagger T &= T && (31)\\ T^\dagger TT^\dagger &= T^\dagger && (32)\\ T^\dagger T &= I-P && (33)\\ T^\dagger T &= Q_{|\mathcal{D}(T^\dagger)} && (34) \end{aligned}$$

hold true.

b)

$$\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp \tag{35}$$

Proof: Because of the definition of $T^\dagger$, for all $y \in \mathcal{D}(T^\dagger)$,

$$T^\dagger y = \tilde{T}^{-1}Qy = T^\dagger Qy,$$

so that $T^\dagger y \in \mathcal{R}(\tilde{T}^{-1}) = \mathcal{N}(T^\dagger)$. For all $x \in \mathcal{N}(T^\dagger)$: $T^\dagger Tx = \tilde{T}^{-1}\tilde{T}x = x$. This proves $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$. Now, for any $y \in \mathcal{D}(T^\dagger)$, (35) implies that $TT^\dagger y = TT^\dagger Qy = T\tilde{T}^{-1}Qy = \tilde{T}\tilde{T}^{-1}Qy = Qy$ since $\tilde{T}^{-1}Qy \in \mathcal{N}(T)^\perp$. Consequently (34) holds. By definition of $T^\dagger$ we have that for all $x \in X$:

$$T^\dagger Tx = \tilde{T}^{-1}T(Px-(I-P)x) = \tilde{T}^{-1}TPx + \tilde{T}^{-1}T(I-P)x = (I-P)x,$$

thus (33) holds. Now (33) implies $TT^\dagger T = T(I-P) = T - TP = T$, therefore (31) is true. (35) and (34) imply (32) ($T^\dagger TT^\dagger = T^\dagger Q_{|D(T^\dagger)} = T^\dagger$). #

Proposition 2 *Let $K : X \to Y$ be compact, $dim\mathcal{R}(T^\dagger) = \infty$. Then $T^\dagger$ is a densly defined unbounded operator.*

We will show that $T^\dagger$ assigns to each $y \in \mathcal{D}(T^\dagger)$ the best approximate solution and that this solution depends discontinuously on y unless $dim\mathcal{R}(T) < \infty$.

Theorem 1 *Let $y \in \mathcal{D}(T^\dagger)$, then*

$$Tx = y, \quad T : X \to Y \tag{36}$$

has a unique best-approximate solution which is given by $T^\dagger y$. The set of all least-squares solutions is $T^\dagger y + \mathcal{N}(T)$.

Proof: Let

$$S = \{z \in X \mid Tz = Qy\}.$$

Since $y \in \mathcal{D}(T^\dagger) = \mathcal{R}(T) + \mathcal{R}(T^\dagger)$, $Qy \in \mathcal{R}(T) \Rightarrow S \neq \emptyset$. As the orthogonal projector Q is also a metric projector, we have $\forall z \in S$ and $\forall x \in X$:

$$||Tz - y|| = ||Qy - y|| \leq ||Tx - y||.$$

So, all elements in S are least-squares of $Tx = y$. Conversely, let z be a least-squares solution of (36). Then

$$||Qy - y|| \leq ||Tz - y|| = \inf\{||u - y|| \mid u \in \mathcal{R}(T)\} = ||Qy - y||.$$

Thus, Tz is the closest element to y in $\mathcal{R}(T)$, i.e. $Tz = Qy$ and

$$S = \{x \in X \mid x \text{ is least-squares solution of } Tx = y\} \neq \emptyset.$$

Now, let $\overline{z}$ be the element of minimal norm in the closed linear manifold $S = T^{-1}(\{Qy\})$. Since then $S = \overline{z} + \mathcal{N}(T)$, it suffices to show that

$$\overline{z} = T^{\dagger}y.$$

As an element of minimal norm in $S = \overline{z} + \mathcal{N}(T)$, $\overline{z}$ is orthogonal to $\mathcal{N}(T)$, i.e. $\overline{z} \in \mathcal{N}(T^{\dagger})$. This implies that $\overline{z} = (I - P)\overline{z} = T^{\dagger}T\overline{z} = T^{\dagger}Qy = T^{\dagger}TT^{\dagger}y = T^{\dagger}y$, i.e. $\overline{z} = T^{\dagger}y$. #

Theorem 2 *Let $y \in \mathcal{D}(T^{\dagger})$. Then $x \in X$ is a least-squares solution of $Tx = y$ if and only if the normal equation*

$$T^*Tx = T^*y \tag{37}$$

holds.

Proof: As we know x is least-squares solution of $Tx = y$ if and only if Tx is the closest element in $\mathcal{R}(T)$ to y, which is equivalent to $Tx - y \in \mathcal{R}(T) = \mathcal{N}(T^*)$, i.e. to $T^*(Tx - y) = 0$ and thus to (37). #

It follows from the last Theorem that $T^{\dagger}y$ is the solution of $T^*Tx = T^*y$ of minimal norm, i.e.

$$T^{\dagger} = (T^*T)^{\dagger}T^*.$$

Theorem 3 *Let $(\sigma_n, u_n, v_n)_{n\in\mathbb{N}}$ be a singular system for K (compact, mapping from X to Y). Then we have*

a)

$$y \in \mathcal{D}(K^{\dagger}) \Longleftrightarrow \sum_{n=1}^{\infty} \frac{|\langle y, v_n\rangle|}{\sigma_n^2} < \infty$$

b) for $y \in \mathcal{D}(K^{\dagger})$

$$K^{\dagger}y = \sum_{n=1}^{\infty} \frac{|\langle y, v_n\rangle|}{\sigma_n} u_n.$$

Proof: Let $y \in \mathcal{D}(K^{\dagger})$, i.e. $Qy \in \mathcal{R}(K)$. The orthogonal projector Q onto $\overline{\mathcal{R}(K)}$ can be written as

$$Q = \sum_{n=1}^{\infty} \langle \cdot, v_n\rangle v_n,$$

since the $\{v_n\}$ span $\overline{\mathcal{R}(K)}$. Since $Qy \in \mathcal{R}(K)$, there exists an $x \in X$ with $Kx = Qy$. W.l.o.g. we can assume that $x \in \mathcal{N}(K)^{\perp}$. Since the $\{u_n\}$ span $\overline{\mathcal{R}(K^{\dagger})} = \mathcal{N}(T)^{\perp}$, $x = \sum_{n=1}^{\infty} \langle x, u_n\rangle u_n$ so that we have

$$\sum_{i=1}^{\infty} \langle y, v_n\rangle v_n = Kx = \sum_{i=1}^{\infty} \langle x, u_n\rangle Ku_n = \sum_{i=1}^{\infty} \sigma_n \langle x, u_n\rangle v_n. \tag{38}$$

Thus, $\forall n \in \mathbb{N}$

$$\langle y, v_n\rangle = \sigma_n \langle x, u_n\rangle$$

must hold. Since as a sequence of Fourier coefficients $(\langle x, u_n \rangle) \in l^2$, so that $\left(\frac{\langle y, v_n \rangle}{\sigma_n}\right) \in l^2$, the condition in *a*) follows. Conversely, assume that this condition holds. By the Riesz - Fischer theorem from functional analysis

$$x := \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} u_n \in X.$$

We have that

$$Kx = \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} K u_n = \sum_{n=1}^{\infty} \langle y, v_n \rangle \, v_n = Qy.$$

Especially, $Qy \in \mathcal{R}(K)$ and hence $y \in \mathcal{D}(K^\dagger)$. Since the $\{u_n\}$ span $\mathcal{N}(K)^\perp$, $x \in \mathcal{N}(K)^\perp$. Since x lies in both, in this set and in $\mathcal{N}(K)^\perp$, x is the element with minimum norm in this set, i.e. $x = K^\dagger Q y = K^\dagger y$. Thus *b*) holds as well. #

References

[1] H.W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.