INVERSE PROBLEME

Vorlesungsskript, Sommersemester 2020

Christian Clason

Stand vom 27. Oktober 2020

Fakultät für Mathematik
Universität Duisburg-Essen

INHALTSVERZEICHNIS

I FUNKTIONALANALYTISCHE GRUNDLAGEN

1 LINEARE OPERATOREN IN NORMIERTEN RÄUMEN 4
- 1.1 Normierte Räume 4
- 1.2 Beschränkte Operatoren 6

2 KOMPAKTE OPERATOREN IN HILBERTRÄUMEN 9
- 2.1 Skalarprodukt und schwache Konvergenz 9
- 2.2 Orthogonalität und Orthogonalsysteme 11
- 2.3 Der Spektralsatz für kompakte Operatoren 12

II LINEARE INVERSE PROBLEME

3 SCHLECHT GESTELLTE OPERATORGLEICHUNGEN 16
- 3.1 Verallgemeinerte Inverse 17
- 3.2 Singulärwertzerlegung kompakter Operatoren 23

4 REGULARISIERUNGSVERFAHREN 31
- 4.1 Regularisierung und Parameterwahl 31
- 4.2 Konvergenzraten 37

5 SPEKTRALE REGULARISIERUNG 45
- 5.1 Regularisierung 47
- 5.2 Parameterwahl und Konvergenzraten 49

6 TIKHONOV-REGULARISIERUNG 58

7 LANDWEBER-REGULARISIERUNG 64

8 DISKRETISIERUNG ALS REGULARISIERUNG 70
- 8.1 Ausgleichsprojektion 70
- 8.2 Duale Ausgleichsprojektion 73

III STATISTISCHE INVERSE PROBLEME

9 FREQUENTISTISCHE INVERSE PROBLEME 79
9.1 Statistisches Fehlermodell und Schätzer 80
9.2 Risiko und Ordnungsoptimalität 82

10 BAYESIANISCHE INVERSE PROBLEME 88

ÜBERBLICK

Inverse Probleme treten überall dort auf, wo sich gesuchte Größen nicht durch direkte Messung ermitteln lassen, sondern nur durch Abgleich von Messungen und mathematischen Modellen. Beispiele finden sich in der biomedizinischen Bildgebung, der zerstörungsfreien Prüfung von Werkstoffen und der Kalibrierung von Finanzmodellen. Der Name „inverses Problem“ rührt daher, dass es oft ein „direktes Problem“ enthält: nämlich jenes mathematische Modell, das die gesuchte Größe mit der Messung verknüpft. Aus mathematischer Sicht ist jedoch relevanter, dass es sich hierbei um sogenannte *schlecht gestellte* Probleme handelt, die mit den Standard-Methoden zur Lösung (nicht)linearer Gleichungen nicht behandelt werden können.[1]

Die mathematische Theorie der inversen Probleme ist daher ein Teilgebiet der Funktionalanalysis: so, wie sich diese mit der Frage beschäftigt, wann eine Gleichung $F(x) = y$ in einem unendlichdimensionalen Vektorraum eine eindeutige Lösung x besitzt, die stetig von y abhängt, werden wir in dieser Vorlesung untersuchen, unter welchen Bedingungen dies nicht der Fall ist, und wie man dann zumindest eine sinnvolle Näherung an x bekommt. Dies entspricht im Wesentlichen dem Schritt von regulären zu inkonsistenten, unterbestimmten, und schlecht konditionierten linearen Gleichungssystemen.

Dieses Skriptum basiert vor allem auf den folgenden Werken:

[1] M. Burger (2007), Inverse Problems, Vorlesungsskript, Institut für Numerische und Angewandte Mathematik, Universität Münster, URL: http://wwwmath.uni-muenster.de/num/Vorlesungen/IP_WS07/skript.pdf

[2] H. W. Engl, M. Hanke & A. Neubauer (1996), *Regularization of Inverse Problems*, Bd. 375, Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, DOI: 10.1007/978-94-009-1740-8

[3] B. von Harrach (2014), Regularisierung Inverser Probleme, Vorlesungsskript, Fachbereich Mathematik, Universität Stuttgart, URL: http://www.mathematik.uni-stuttgart.de/~harrach/lehre/Regularisierung.pdf

[1] Sonst wäre auch keine eigene Vorlesung notwendig. Tatsächlich wäre der Titel „Schlecht gestellte Probleme“ passender; die Bezeichnung „Inverse Probleme“ hat sich dennoch durchgesetzt.

[4] T. Hohage (2002), Inverse Problems, Vorlesungsskript, Institut für Numerische und Angewandte Mathematik, Universität Göttingen

[5] K. Ito & B. Jin (2014), *Inverse Problems: Tikhonov Theory and Algorithms*, Bd. 22, Series on Applied Mathematics, World Scientific, Singapore, DOI: 10.1142/9789814596206_0001

[6] A. Kirsch (2011), *An Introduction to the Mathematical Theory of Inverse Problems*, 2. Aufl., Springer, New York, DOI: 10.1007/978-1-4419-8474-6

[7] A. K. Louis (1989), *Inverse und schlecht gestellte Probleme*, Teubner Studienbücher Mathematik, B. G. Teubner, Stuttgart, DOI: 10.1007/978-3-322-84808-6

[8] A. Rieder (2003), *Keine Probleme mit inversen Problemen*, Eine Einführung in ihre stabile Lösung, Friedr. Vieweg & Sohn, Braunschweig, DOI: 10.1007/978-3-322-80234-7

[9] L. Cavalier (2011), Inverse problems in statistics, in: *Inverse Problems and High-Dimensional Estimation*, Bd. 203, Lect. Notes Stat. Proc. Springer, Heidelberg, 3–96, DOI: 10.1007/978-3-642-19989-9_1

[10] H. Kekkonnen (2019), Bayesian Inverse Problems, Vorlesungsskript, University of Cambridge, URL: http://www.damtp.cam.ac.uk/research/cia/files/teaching/Inverse_Problems_Lent_2018/19y_01m_16d_LectureNotes.pdf

Teil I

FUNKTIONALANALYTISCHE GRUNDLAGEN

1 LINEARE OPERATOREN IN NORMIERTEN RÄUMEN

In diesem und dem nächsten Kapitel stellen wir die für diese Vorlesung wesentlichen Begriffe, Notationen und Resultate zusammen. Für Beweise wird auf die Standardliteratur verwiesen, z. B. auf [Alt 2012; Werner 2011], oder auf [Clason 2019].

1.1 NORMIERTE RÄUME

Im Folgenden bezeichne X einen Vektorraum über dem Körper $\mathbb{K}$, wobei wir uns hier stets auf den Fall $\mathbb{K} = \mathbb{R}$ beschränken. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}^+ := [0, \infty)$ heißt *Norm* (auf X), falls für alle $x \in X$ gilt

(i) $\|\lambda x\| = |\lambda| \|x\|$ für alle $\lambda \in \mathbb{K}$,

(ii) $\|x + y\| \leq \|x\| + \|y\|$ für alle $y \in X$,

(iii) $\|x\| = 0$ genau dann, wenn $x = 0 \in X$.

Zwei Normen $\|\cdot\|_1$, $\|\cdot\|_2$ heißen *äquivalent*, falls $c_1, c_2 > 0$ existieren mit

$$c_1 \|x\|_2 \leq \|x\|_1 \leq c_2 \|x\|_2 \qquad \text{für alle } x \in X.$$

Ist X endlichdimensional, so sind alle Normen auf X äquivalent. Die Konstanten c_1, c_2 hängen dann jedoch von der Dimension von X ab; insbesondere können $c_1(N) \to 0$ oder $c_2(N) \to \infty$ für $\dim X = N \to \infty$ gehen, was die entsprechende Ungleichung für wachsende Dimensionen wertlos macht. Die Vermeidung solcher dimensionsabhängiger Konstanten ist daher einer der Gründe, warum wir inverse Probleme in unendlichdimensionalen Räumen betrachten.

Beispiel 1.1. (i) Auf $X = \mathbb{R}^N$ werden Normen definiert durch

$$\|x\|_p = \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p}, \qquad 1 \leq p < \infty,$$
$$\|x\|_\infty = \max_{i=1,\dots,N} |x_i|.$$

(ii) Auf $X = \ell^p$ (dem Raum der reellen Folgen, auf dem folgende Ausdrücke endlich sind) sind Normen definiert durch

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}, \qquad 1 \le p < \infty,$$
$$\|x\|_\infty = \sup_{i=1,\dots,\infty} |x_i|.$$

(iii) Auf $X = L^p(\Omega)$ (dem Raum der messbaren reellen Funktionen, auf dem folgende Ausdrücke endlich sind) sind Normen definiert durch

$$\|u\|_p = \left(\int_\Omega |u(x)|^p \, dx\right)^{1/p}, \qquad 1 \le p < \infty,$$
$$\|u\|_\infty = \operatorname*{ess\,sup}_{x\in\Omega} |u(x)|.$$

(iv) Auf $X = C(\overline{\Omega})$ (dem Raum der Funktionen auf Ω, die stetig auf den Rand fortgesetzt werden können) ist eine Norm definiert durch

$$\|u\|_C = \sup_{x\in\Omega} |u(x)|.$$

Auf ähnliche Weise definiert man den Raum $C^k(\overline{\Omega})$ der k-mal stetig differenzierbaren Funktionen mit Norm $\|u\|_{C^k} = \sum_{j=0}^{k} \|u^{(j)}\|_C$.

Ist $\|\cdot\|$ eine Norm auf X, so bezeichnet man das Paar $(X, \|\cdot\|)$ als *normierten Raum*, und schreibt in diesem Fall oft $\|\cdot\|_X$. Ist die Norm kanonisch (etwa in Beispiel 1.1 (ii)–(iv)), so wird sie oft weggelassen.

Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume mit $X \subset Y$, so heißt X *stetig eingebettet* in Y, geschrieben $X \hookrightarrow Y$, falls ein $C > 0$ existiert mit

$$\|x\|_Y \le C\|x\|_X \qquad \text{für alle } x \in X.$$

Eine Norm vermittelt auf direkte Weise einen Konvergenzbegriff, die sogenannte *starke Konvergenz*: Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ konvergiert (stark in X) gegen ein $x \in X$, geschrieben $x_n \to x$, wenn gilt

$$\lim_{n\to\infty} \|x_n - x\|_X = 0.$$

Eine Teilmenge $U \subset X$ nennen wir

- *abgeschlossen*, falls für jede konvergente Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ auch der Grenzwert $x \in U$ liegt;

- *kompakt*, falls jede Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ eine konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ mit Grenzwert $x \in U$ besitzt;
- *dicht* in X, falls für alle $x \in X$ eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset U$ mit $x_n \to x$ existiert.

Die Vereinigung von U mit der Menge aller Grenzwerte von konvergenten Folgen in U bezeichnet man als ihren *Abschluss* $\overline{U}$; offensichtlich ist U dicht in $\overline{U}$.

Ein normierter Raum X heißt *vollständig*, wenn jede Cauchy-Folge in X konvergiert; man nennt dann X auch *Banachraum*. Alle Räume in Beispiel 1.1 sind Banachräume. Ist X ein unvollständiger normierter Raum, so bezeichnet man $\overline{X}$ als *Vervollständigung* von X (bezüglich der Norm $\|\cdot\|_X$).

Schließlich definieren wir für späteren Gebrauch für $x \in X$ und $r > 0$

- die *offene Kugel* $U_r(x) := \{z \in X : \|x - z\|_X < r\}$ und
- die *abgeschlossene Kugel* $B_r(x) := \{z \in X : \|x - z\|_X \leq r\}$.

Die abgeschlossene Kugel um 0 mit Radius 1 bezeichnet man auch als *Einheitskugel* B_X. Eine Menge $U \subset X$ heißt

- *offen*, falls für alle $x \in U$ ein $r > 0$ existiert mit $U_r(x) \subset U$ (d. h. alle $x \in U$ *innere Punkte* von U sind);
- *beschränkt*, falls sie in einer abgeschlossenen Kugel $B_r(0)$ für ein $r > 0$ enthalten ist;
- *konvex*, falls für $x, y \in U$ auch $\lambda x + (1-\lambda)y \in U$ für alle $\lambda \in [0,1]$ gilt.

In normierten Räumen gilt, dass das Komplement einer offenen Menge abgeschlossen ist und umgekehrt (d. h. die abgeschlossenen Mengen im Sinne der Topologie sind genau die (Folgen-)abgeschlossenen Mengen im Sinne unserer Definition). Sowohl offene als auch abgeschlossene Kugeln sind wegen der Norm-Axiome konvex. Dagegen ist die Einheitskugel dann *und nur dann* kompakt, wenn X endlichdimensional ist; dies wird wesentliche Auswirkungen für uns haben.

1.2 BESCHRÄNKTE OPERATOREN

Wir betrachten nun Abbildungen zwischen normierten Räumen. Seien im Folgenden stets $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume, $U \subset X$, und $F : U \to Y$ eine Abbildung. Wir bezeichnen mit

- $\mathcal{D}(F) := U$ den *Definitionsbereich* (englisch „domain“) von F;
- $\mathcal{N}(F) := \{x \in U : F(x) = 0\}$ den *Kern* (englisch „kernel“ oder „null space“) von F;
- $\mathcal{R}(F) := \{F(x) \in Y : x \in U\}$ das *Bild* (englisch „range“) von F.

Wir sagen, F ist

- *stetig* in $x \in U$, wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\|F(x) - F(z)\|_Y \leq \varepsilon \qquad \text{für alle } z \in U \text{ mit } \|x - z\|_X \leq \delta;$$

- *Lipschitz-stetig*, wenn ein $L > 0$ existiert (genannt *Lipschitz-Konstante*) mit

$$\|F(x_1) - F(x_2)\|_Y \leq L\|x_1 - x_2\|_X \qquad \text{für alle } x_1, x_2 \in U.$$

Eine Abbildung $F : X \to Y$ ist also genau dann stetig, wenn aus $x_n \to x$ auch $F(x_n) \to F(x)$ folgt; sie ist *abgeschlossen*, wenn für $x_n \to x$ und $F(x_n) \to y$ folgt, dass $F(x) = y$ ist.

Ist $F : X \to Y$ linear (d. h. $F(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 F(x_1) + \lambda_2 F(x_2)$ für alle $\lambda_1, \lambda_2 \in \mathbb{R}$ und $x_1, x_2 \in X$), so ist die Stetigkeit äquivalent mit der Bedingung, dass eine Konstante $C > 0$ existiert mit

$$\|F(x)\|_Y \leq C\|x\|_X \qquad \text{für alle } x \in X.$$

Stetige lineare Abbildungen nennt man daher auch *beschränkt*; man spricht auch von einem beschränkten linearen *Operator*. (Diese bezeichnen wir in Folge mit T und lassen die Klammern um das Argument weg, um dies zu verdeutlichen.) Ist Y vollständig, so ist der Raum $\mathcal{L}(X, Y)$ der beschränkten linearen Operatoren ein Banachraum versehen mit der *Operatornorm*

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{x \in X \setminus \{0\}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y = \sup_{\|x\|_X = 1} \|Tx\|_Y,$$

die gleich der minimalen Konstante C in der Bedingung für die Stetigkeit ist. Daraus folgt auch sofort

$$\|Tx\|_Y \leq \|T\|_{\mathcal{L}(X,Y)}\|x\|_X \qquad \text{für alle } x \in X.$$

Wie in der linearen Algebra bezeichnet man T als

- *injektiv*, wenn $\mathcal{N}(T) = \{0\}$ ist;
- *surjektiv*, wenn $\mathcal{R}(T) = Y$ ist;
- *bijektiv*, wenn T injektiv und surjektiv ist.

Ist $T \in \mathcal{L}(X, Y)$ bijektiv, dann ist die Inverse $T^{-1} : Y \to X$ stetig genau dann, wenn ein $c > 0$ existiert mit

$$c\|x\|_X \leq \|Tx\|_Y \qquad \text{für alle } x \in X; \tag{1.1}$$

in diesem Fall gilt $\|T^{-1}\|_{\mathcal{L}(Y,X)} = c^{-1}$ für das größte solche c. Wann dies der Fall ist, besagen die folgenden Hauptsätze der Funktionalanalysis (die alle mehr oder weniger direkt aus dem Satz von der offenen Abbildung folgen).

Satz 1.2 (von der stetigen Inversen). *Seien X, Y Banachräume und $T \in \mathcal{L}(X, Y)$ bijektiv. Dann ist $T^{-1} : Y \to X$ stetig.*

Für inverse Probleme besonders relevant ist der Fall, dass T zwar injektiv, aber nicht surjektiv ist. Dann hätte man gerne zumindest auf dem Bild von T eine stetige Inverse; dass dies im Allgemeinen nicht mehr der Fall ist, stellt die Kernschwierigkeit bei unendlichdimensionalen inversen Problemen dar.

Satz 1.3 (vom abgeschlossenen Bild). *Seien X, Y Banachräume und $T \in \mathcal{L}(X, Y)$ injektiv. Dann ist $T^{-1} : \mathcal{R}(T) \to X$ stetig genau dann, wenn $\mathcal{R}(T)$ abgeschlossen ist.*

Das Trio komplett macht der folgende Satz.

Satz 1.4 (vom abgeschlossenen Graphen). *Seien X, Y Banachräume. Dann ist $T : X \to Y$ stetig genau dann, wenn T abgeschlossen ist.*

Wir betrachten nun Folgen linearer Operatoren. Dafür unterscheiden wir zwei Konvergenzbegriffe: Eine Folge $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{L}(X, Y)$ konvergiert gegen $T \in \mathcal{L}(X, Y)$

(i) *punktweise*, wenn $T_n x \to Tx$ (stark in Y) für alle $x \in X$ konvergiert;

(ii) *gleichmäßig*, wenn $T_n \to T$ (stark in $\mathcal{L}(X, Y)$) konvergiert.

Gleichmäßige Konvergenz impliziert dabei punktweise Konvergenz; schwächere Bedingungen liefert ein weiterer Hauptsatz der Funktionalanalysis.

Satz 1.5 (Banach–Steinhaus). *Sei X ein Banachraum, Y ein normierter Vektorraum, und $\{T_i\}_{i\in I} \subset \mathcal{L}(X, Y)$ eine Familie punktweise beschränkter Operatoren, d. h. für alle $x \in X$ existiert eine Konstante $M_x > 0$ mit $\sup_{i\in I} \|T_i x\|_Y \leq M_x$. Dann ist*

$$\sup_{i\in I} \|T_i\|_{\mathcal{L}(X,Y)} < \infty.$$

Folgerung 1.6. *Seien X, Y Banachräume und $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{L}(X, Y)$. Dann sind äquivalent*

(i) $\{T_n\}_{n\in\mathbb{N}}$ konvergiert gleichmäßig auf kompakten Teilmengen in X;

(ii) $\{T_n\}_{n\in\mathbb{N}}$ konvergiert punktweise auf X;

(iii) $\{T_n\}_{n\in\mathbb{N}}$ konvergiert punktweise auf einer dichten Teilmenge $U \subset X$ und es gilt

$$\sup_{n\in\mathbb{N}} \|T_n\|_{\mathcal{L}(X,Y)} < \infty.$$

Folgerung 1.7. *Seien X, Y Banachräume und $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{L}(X, Y)$. Konvergiert T_n punktweise gegen ein $T : X \to Y$, so ist T beschränkt.*

2 KOMPAKTE OPERATOREN IN HILBERTRÄUMEN

Zwar kann man inverse Probleme in Banachräumen untersuchen, in Hilberträumen lässt sich die Theorie für lineare Operatoren aber weitestgehend geschlossen darstellen. Dort wird auch die Analogie zu unterbestimmten und schlecht konditionierten Gleichungssystemen besonders deutlich.

2.1 SKALARPRODUKT UND SCHWACHE KONVERGENZ

Hilberträume zeichnen sich dadurch aus, dass eine zusätzliche Struktur definiert ist: Eine Abbildung $(\cdot,\cdot) : X \times X \to \mathbb{R}$ auf dem Vektorraum X über $\mathbb{R}$ heißt *Skalarprodukt*, falls gilt

(i) $(\alpha x + \beta y, z) = \alpha\,(x,z) + \beta\,(y,z)$ für alle $x, y, z \in X$ und $\alpha, \beta \in \mathbb{R}$;

(ii) $(x,y) = (y,x)$ für alle $x, y \in X$;

(iii) $(x,x) \geq 0$ für alle $x \in X$ mit $(x,x) = 0$ genau dann, wenn $x = 0$ ist.

Durch das Skalarprodukt wird eine Norm

$$\|x\|_X := \sqrt{(x,x)_X}$$

induziert, die der *Cauchy–Schwarz-Ungleichung* gehorcht:

$$|\,(x,y)_X\,| \leq \|x\|_X \|y\|_X.$$

(Für festgehaltenes y ist das Skalarprodukt also stetig in x – bezüglich der induzierten Norm – und umgekehrt.) Ist X bezüglich der induzierten Norm vollständig (also $(X, \|\cdot\|_X)$ ein Banachraum), so wird X *Hilbertraum* genannt; ist das Skalarprodukt und damit die induzierte Norm kanonisch, lässt man es weg. Beispiel 1.1 (i–iii) für $p = 2$ ist jeweils ein Hilbertraum, wobei das Skalarprodukt gegeben ist durch

(i) für $X = \mathbb{R}^N$: $\quad (x,y)_X = \sum_{i=1}^{N} x_i y_i,$

(ii) für $X = \ell^2$: $\quad (x,y)_X = \sum_{i=1}^{\infty} x_i y_i,$

(iii) für $X = L^2(\Omega)$: $\quad (u, v)_X = \int_\Omega u(x)v(x)\,dx.$

In allen Fällen induziert das Skalarprodukt die kanonische Norm.

Durch das Skalarprodukt wird ein weiterer Konvergenzbegriff erzeugt: die *schwache Konvergenz*. Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ konvergiert schwach (in X) gegen $x \in X$, geschrieben $x_n \rightharpoonup x$, falls gilt

$$(x_n, z)_X \to (x, z)_X \qquad \text{für alle } z \in X.$$

Dieser Begriff verallgemeinert die komponentenweise Konvergenz im $\mathbb{R}^N$ (wähle $z = e_i$, den i-ten Einheitsvektor); in endlichdimensionalen Räumen fallen also starke und schwache Konvergenz zusammen. In unendlichdimensionalen Räumen impliziert starke Konvergenz die schwache, aber nicht umgekehrt. Konvergiert eine Folge $\{x_n\}_{n\in\mathbb{N}}$ jedoch schwach gegen ein $x \in X$, und gilt zusätzlich $\|x_n\|_X \to \|x\|_X$, so konvergiert x_n auch stark gegen x. Es gilt außerdem, dass die Norm *schwach unterhalbstetig* ist: Konvergiert $x_n \rightharpoonup x$, so ist

$$\|x\|_X \le \liminf_{n\to\infty} \|x_n\|_X. \tag{2.1}$$

Dieser Konvergenzbegriff ist vor allem deshalb nützlich, weil für ihn – im Gegensatz zur starken Konvergenz! – auch im Unendlichdimensionalen der Satz von Bolzano–Weierstrass gilt: Jede beschränkte Folge in einem Hilbertraum besitzt eine schwach konvergente Teilfolge. Umgekehrt ist jede schwach konvergente Folge beschränkt.

Wir betrachten nun lineare Operatoren $T \in \mathcal{L}(X, Y)$ zwischen Hilberträumen X, Y. Von besonderer Bedeutung ist dabei der Spezialfall $Y = \mathbb{R}$ sein, das heißt der Raum $\mathcal{L}(X, \mathbb{R})$ der *linearen stetigen Funktionale* auf X. Diese lassen sich mit Elementen von X identifizieren.

Satz 2.1 (Fréchet–Riesz). *Sei X ein Hilbertraum. Dann existiert zu jedem $\lambda \in \mathcal{L}(X, \mathbb{R})$ genau ein $z_\lambda \in X$ mit $\|\lambda\|_{\mathcal{L}(X,\mathbb{R})} = \|z_\lambda\|_X$ und*

$$\lambda(x) = (z_\lambda, x)_X \qquad \textit{für alle } x \in X.$$

Mit Hilfe dieses Satzes kann für jeden linearen Operator $T \in \mathcal{L}(X, Y)$ ein *adjungierter Operator* $T^* \in \mathcal{L}(Y, X)$ definiert werden durch

$$(T^* y, x)_X = (Tx, y)_Y \qquad \text{für alle } x \in X,\ y \in Y,$$

für den gilt

(i) $(T^*)^* = T$;

(ii) $\|T^*\|_{\mathcal{L}(Y,X)} = \|T\|_{\mathcal{L}(X,Y)}$;

(iii) $\|T^*T\|_{\mathcal{L}(X,X)} = \|T\|^2_{\mathcal{L}(X,Y)}$.

Ist $T^* = T$, so nennt man T *selbstadjungiert*.

2.2 ORTHOGONALITÄT UND ORTHOGONALSYSTEME

Ein Skalarprodukt vermittelt den Begriff der *Orthogonalität*: Ist X ein Hilbertraum, so nennt man $x, y \in X$ *orthogonal*, falls $(x, y)_X = 0$ gilt. Für eine Teilmenge $U \subset X$ heißt

$$U^\perp := \{x \in X : (x, u)_X = 0 \text{ für alle } u \in U\}$$

das *orthogonale Komplement* von U in X; direkt aus der Definition folgt, dass $U^\perp$ ein abgeschlossener Unterraum ist. Insbesondere ist $X^\perp = \{0\}$.

Weiterhin folgt, dass $U \subset (U^\perp)^\perp$ ist. Ist U ein abgeschlossener Unterraum, so gilt sogar $U = (U^\perp)^\perp$ (und damit $\{0\}^\perp = X$). In diesem Fall existiert die *Orthogonalzerlegung*

$$X = U \oplus U^\perp,$$

d. h. jedes Element $x \in X$ kann auf eindeutige Weise dargestellt werden als

$$x = u + u_\perp, \qquad u \in U,\ u_\perp \in U^\perp.$$

Durch die Zuordnung $x \mapsto u$ wird ein linearer Operator $P_U \in \mathcal{L}(X, X)$ definiert, den man *Orthogonalprojektion* auf U nennt. Für diesen gilt:

(i) P_U ist selbstadjungiert;

(ii) $\|P_U\|_{\mathcal{L}(X,X)} = 1$;

(iii) $\mathrm{Id} - P_U = P_{U^\perp}$;

(iv) $\|x - P_U x\|_X = \min_{u \in U} \|x - u\|_X$;

(v) $z = P_U x$ genau dann, wenn $z \in U$ und $z - x \in U^\perp$.

Ist der Unterraum U dagegen nicht abgeschlossen, gilt lediglich $(U^\perp)^\perp = \overline{U} \supset U$. Für $T \in \mathcal{L}(X, Y)$ gilt daher

(i) $\mathcal{R}(T)^\perp = \mathcal{N}(T^*)$ und damit $\mathcal{N}(T^*)^\perp = \overline{\mathcal{R}(T)}$;

(ii) $\mathcal{R}(T^*)^\perp = \mathcal{N}(T)$ und damit $\mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$.

Insbesondere ist der Kern einer stetigen linearen Abbildung stets abgeschlossen; außerdem ist T injektiv genau dann, wenn $\mathcal{R}(T^*)$ dicht in X liegt.

Eine Menge $U \subset X$, deren Elemente paarweise orthogonal sind, heißt *Orthogonalsystem*. Gilt sogar für alle $x, y \in U$, dass

$$(x, y)_X = \begin{cases} 1 & \text{falls } x = y, \\ 0 & \text{sonst,} \end{cases}$$

so heißt U *Orthonormalsystem*. Ein Orthonormalsystem ist *vollständig*, falls kein Orthonormalsystem $V \subset X$ mit $U \subsetneq V$ existiert. Jedes Orthonormalsystem $U \subset X$ erfüllt die *Besselsche Ungleichung*:

$$\sum_{u \in U} |\,(x,u)_X\,|^2 \leq \|x\|_X^2 \qquad \text{für alle } x \in X, \tag{2.2}$$

wobei stets höchstens abzählbar viele Summanden von Null verschieden sind. Gilt sogar Gleichheit in (2.2), so nennt man U auch *Orthonormalbasis*; in diesem Fall ist U vollständig und es gilt

$$x = \sum_{u \in U} (x,u)_X\, u \qquad \text{für alle } x \in X. \tag{2.3}$$

Jeder Hilbertraum besitzt eine Orthonormalbasis. Ist diese höchstens abzählbar, so nennt man den Hilbertraum *separabel*. Aus der Besselschen Ungleichung folgt dann, dass die Folge $\{u_n\}_{n\in\mathbb{N}} = U$ schwach gegen Null (aber wegen $\|u_n\|_X = 1$ nicht stark!) konvergiert. Ein Beispiel ist $X = L^2([0,1])$, welches die Orthonormalbasis $\{u_n\}_{n\in\mathbb{Z}}$ mit

$$u_n(x) = \begin{cases} \sqrt{2}\sin(2\pi\, n\, x) & n > 0, \\ \sqrt{2}\cos(2\pi\, n\, x) & n < 0, \\ 1 & n = 0. \end{cases}$$

besitzt.

Schließlich halten wir fest, dass jeder abgeschlossene Unterraum $U \subset X$ eine Orthonormalbasis $\{u_n\}_{n\in\mathbb{N}}$ besitzt, mit deren Hilfe die Orthogonalprojektion auf U dargestellt werden kann als

$$P_U x = \sum_{j=1}^{\infty} \big(x, u_j\big)_X\, u_j.$$

2.3 DER SPEKTRALSATZ FÜR KOMPAKTE OPERATOREN

So wie man Hilberträume als naheliegende Verallgemeinerung von endlichdimensionalen Vektorräumen auffassen kann, sind kompakte Operatoren das unendlichdimensionale Analogon zu Matrizen. Ein linearer Operator $T : X \to Y$ heißt dabei *kompakt*, wenn das Bild jeder beschränkten Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ eine konvergente Teilfolge $\{Tx_{n_k}\}_{k\in\mathbb{N}} \subset Y$ enthält. Eine äquivalente Charakterisierung ist die folgende: Ein linearer Operator T ist kompakt genau dann, wenn T schwach konvergente Folgen in X auf stark konvergente Folgen in Y abbildet. (Diese Eigenschaft wird auch *Vollstetigkeit* genannt.) Kompakte Operatoren bezeichnen wir in der Regel mit K.

Offensichtlich ist jeder lineare Operator mit endlichdimensionalem Bild kompakt. Insbesondere ist die Identität $\mathrm{Id} : X \to X$ genau wie die Einheitskugel B_X kompakt dann *und*

nur dann, wenn X endlichdimensional ist. Weiterhin gilt, dass der Raum $\mathcal{K}(X, Y)$ aller linearen kompakten Operatoren von X nach Y einen abgeschlossenen Unterraum von $\mathcal{L}(X, Y)$ (und damit einen Banachraum, versehen mit der Operatornorm) bildet. Also ist auch jeder Grenzwert von linearen Operatoren mit endlichdimensionalem Bild kompakt. Sind $T \in \mathcal{L}(X, Y)$ und $S \in \mathcal{L}(Y, Z)$, und ist wenigstens einer der beiden Operatoren kompakt, so ist auch $S \circ T$ kompakt. Weiterhin ist T^* kompakt genau dann, wenn T kompakt ist (Satz von Schauder).

Ein kanonisches Beispiel für kompakte Operatoren sind *Integraloperatoren*. Wir betrachten $X = Y = L^2(\Omega)$ mit $\Omega = (0, 1)$ und für einen gegebenen *Kern* $k \in L^2(\Omega \times \Omega)$ den Operator $K : L^2(\Omega) \to L^2(\Omega)$, der punktweise definiert ist durch

$$[Kx](t) = \int_0^1 k(s, t)x(s)\, ds \qquad \text{für fast alle } t \in \Omega$$

(wobei $Kx \in L^2(\Omega)$ nach dem Satz von Fubini gilt). Mit Hilfe der Cauchy–Schwarz-Ungleichung und dem Satz von Fubini erhält man nun sofort

$$\|K\|_{\mathcal{L}(X,Y)} \leq \|k\|_{L^2(\Omega^2)},$$

woraus auch folgt, dass K ein beschränkter Operator von $L^2(\Omega)$ nach $L^2(\Omega)$ ist.

Da $k \in L^2(\Omega^2)$ messbar ist, existiert eine Folge $\{k_n\}_{n\in\mathbb{N}}$ von einfachen Funktionen (d. h. solche, die nur endlich viele verschiedene Werte annehmen) mit $k_n \to k$ in $L^2(\Omega^2)$. Diese haben z. B. die Form

$$k_n(s, t) = \sum_{i,j=1}^{n} \alpha_{ij} \chi_{E_i}(s) \chi_{E_j}(t),$$

wobei χ_E die charakteristische Funktion des messbaren Intervalls $E \subset \Omega$ ist und die E_i eine endliche disjunkte Zerlegung von Ω bilden. Also gilt für den entsprechenden Integraloperator K_n mit Kern k_n wegen der Linearität des Integrals

$$\|K_n - K\|_{\mathcal{L}(X,Y)} \leq \|k_n - k\|_{L^2(\Omega^2)} \to 0,$$

d. h. $K_n \to K$. Nun ist aber

$$[K_n x](t) = \int_0^1 k_n(s, t)x(s)\, ds = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha_{ij} \int_{E_i} x(s)\, ds \right) \chi_{E_j}(t)$$

und damit $K_n x$ eine Linearkombination der $\{\chi_{E_j}\}_{1\leq j\leq n}$. Also ist K Grenzwert einer Folge $\{K_n\}_{n\in\mathbb{N}}$ von Operatoren mit endlichdimensionalem Bild und daher kompakt.

Für den adjungierten Operator $K^* \in \mathcal{L}(X, X)$ rechnet man mit dem Satz von Fubini nach, dass gilt

$$[K^* y](s) = \int_0^1 k(s, t)\, y(t)\, dt \qquad \text{für fast alle } s \in \Omega.$$

Ein Integraloperator ist also genau dann selbstadjungiert, wenn der Kern symmetrisch ist, d. h. $k(s,t) = k(t,s)$ für (fast) alle $s, t \in \Omega$ gilt.

Durch geeignete Wahl des Kerns sieht man, dass zum Beispiel Lösungsoperatoren für partielle Differentialgleichungen oder Faltungsoperatoren – und damit eine große Klasse von praktisch relevanten Operatoren – als Integraloperatoren dargestellt werden können und daher kompakt sind.

Die Analogie zwischen kompakten Operatoren und Matrizen besteht vor allem darin, dass kompakte lineare Operatoren höchstens abzählbar viele Eigenwerte besitzen. (Dies ist für beschränkte lineare Operatoren nicht notwendigerweise der Fall!) Es gilt auch die folgende Variante der Schur-Faktorisierung; sie ist wesentlich dafür verantwortlich, dass inverse Probleme in Hilberträumen so gründlich untersucht werden können.

Satz 2.2 (Spektralsatz). *Sei X ein Hilbertraum und sei $K \in \mathcal{K}(X,X)$ selbstadjungiert. Dann existiert ein (möglicherweise endliches) Orthonormalsystem $\{u_n\}_{n\in\mathbb{N}} \subset X$ und eine (in diesem Fall abbrechende) Nullfolge $\{\lambda_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ mit*

$$Kx = \sum_{n\in\mathbb{N}} \lambda_n \, (x, u_n)_X \, u_n \qquad \text{für alle } x \in X.$$

Weiterhin bilden die $\{u_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K)}$.

Setzt man $x = u_n$, so folgt sofort, dass u_n Eigenvektor zum Eigenwert λ_n ist, d. h. $Ku_n = \lambda_n u_n$ gilt. Üblicherweise ordnet man die Eigenwerte nach abfallendem Betrag, d. h.

$$|\lambda_1| \geq |\lambda_2| \geq \cdots > 0.$$

Dann ist $\|K\|_{L(X,X)} = |\lambda_1|$.

Teil II

LINEARE INVERSE PROBLEME

3 SCHLECHT GESTELLTE OPERATORGLEICHUNGEN

Wir beginnen nun unsere Untersuchung von Operatorgleichungen, die nicht mit Standardmethoden lösbar sind. Wir betrachten zuerst einen linearen Operator T zwischen zwei normierten Räumen X und Y. Nach Hadamard nennen wir die Gleichung $Tx = y$ *korrekt gestellt*, wenn für alle $y \in Y$

(i) ein $x \in X$ existiert mit $Tx = y$;

(ii) diese Lösung eindeutig ist, d. h. $z \neq x$ impliziert $Tz \neq y$;

(iii) diese Lösung stetig von y abhängt, d. h. für alle $\{x_n\}_{n\in\mathbb{N}}$ mit $Tx_n \to y$ gilt $x_n \to x$.

Ist eine dieser Bedingungen verletzt, so nennen wir die Gleichung *schlecht gestellt*.

Eine Verletzung der ersten beiden Bedingungen rührt in der Praxis oftmals von ungenügendem Wissen über die Realität her, und kann durch Erweiterung des zugrundeliegenden mathematischen Modells behandelt werden. Zum Beispiel kann das Konzept einer Lösung erweitert werden, so dass für beliebige $y \in Y$ eine verallgemeinerte Lösung existiert. Ist diese nicht eindeutig, so kann anhand von Zusatzinformationen über das gesuchte x eine bestimmte Lösung ausgewählt werden. Für endlichdimensionale Hilberträume führt dies auf das bekannte Prinzip der *Ausgleichsrechnung*; da dort alle linearen Operatoren stetig sind, ist damit das Problem im Prinzip gelöst (auch wenn die Details und insbesondere die effiziente numerische Lösung noch viel Arbeit erfordern). In unendlichdimensionalen Räumen ist dies jedoch nicht der Fall, wie das folgende Beispiel illustrieren soll.

Beispiel 3.1. Wir suchen zu gegebenem $y \in Y := C^1([0,1])$ die Ableitung $x := y' \in C([0,1])$. Wir nehmen nun an, dass die abzuleitende Funktion y nur anhand von Messwerten gegeben ist, die mit additiven Fehlern behaftet sind, d. h. wir haben nur

$$\tilde{y} = \bar{y} + \eta$$

zur Verfügung. In der Regel können wir nicht davon ausgehen, dass die Störung η ebenfalls stetig differenzierbar ist; nehmen wir aber der Einfachheit an, dass sie zumindest stetig ist. Dann ist $\tilde{y}$ ebenfalls in $C([0,1])$, und wir müssen die Abbildung $x = y' \mapsto y$ als (linearen) Operator $T : C([0,1]) \to C([0,1])$ auffassen. Offensichtlich ist damit Bedingung (i) verletzt.

Aber selbst im Fall, dass die Störung zufällig stetig differenzierbar ist, ist das Problem nicht korrekt gestellt: Betrachte eine Folge $\{\delta_n\}_{n\in\mathbb{N}}$ mit $\delta_n \to 0$, wähle $k \in \mathbb{N}$ beliebig und setze

$$\eta_n(t) := \delta_n \sin\left(\frac{kt}{\delta_n}\right)$$

sowie $\tilde{y}_n := y + \eta_n$. Dann ist $\eta_n \in C^1([0,1])$ und

$$\|\tilde{y}_n - y\|_C = \|\eta_n\|_C = \delta_n \to 0$$

aber

$$\tilde{x}_n(t) := \tilde{y}_n'(t) = y'(t) + k\cos\left(\frac{kt}{\delta_n}\right),$$

d. h. für $x := y'$ gilt

$$\|x - \tilde{x}_n\|_C = \|\eta_n'\|_C = k \qquad \text{für alle } n \in \mathbb{N}.$$

Der Fehler in der Ableitung x kann also (in Abhängigkeit von k) beliebig groß sein, auch wenn der Fehler in y beliebig klein ist.

(Dagegen ist das Problem mit $T : C([0,1]) \to C^1([0,1])$ natürlich korrekt gestellt, denn dann impliziert $\|\eta_n\|_{C^1} \to 0$ nach Definition $\|\bar{x}-x_n\|_C \le \|\eta_n\|_{C^1} \to 0$. Die zugrundeliegende Norm entscheidet also wesentlich über die Schlechtgestelltheit; diese ist allerdings in der Regel durch die Problemstellung vorgegeben – in unserem Beispiel bedeutet die Wahl von $C^1([0,1])$ als Bildraum, dass neben y auch y' gemessen wird, was ja genau die zu bestimmende Größe ist.)

Beachten Sie auch, dass die drei Bedingungen für die Korrektgestelltheit nicht völlig unabhängig sind. Erfüllt $T \in \mathcal{L}(X,Y)$ zum Beispiel die ersten beiden Bedingungen und sind X und Y Banachräume, so ist T bijektiv und hat daher nach Satz 1.2 eine stetige Inverse, womit auch automatisch die dritte Bedingung erfüllt ist.

3.1 VERALLGEMEINERTE INVERSE

Wir versuchen nun, die ersten beiden Bedingungen für lineare Operatoren zwischen Hilberträumen zu erfüllen, indem wir das Konzept der Lösung analog zur Ausgleichsrechnung im $\mathbb{R}^N$ verallgemeinern. Seien X, Y Hilberträume (was wir von nun an voraussetzen) und betrachte für $T \in \mathcal{L}(X,Y)$ die Gleichung $Tx = y$. Ist $y \notin \mathcal{R}(T)$, so hat die Gleichung keine Lösung. In diesem Fall ist es sinnvoll, ein $x \in X$ zu finden, das den Abstand $\|Tx - y\|_Y$ minimiert. Ist andererseits $\mathcal{N}(T) \neq \{0\}$, so existieren unendlich viele Lösungen; wir wählen in diesem Fall diejenige aus, die minimale Norm $\|x\|_X$ hat. Dies führt auf die folgende Definition.

Definition 3.2. Ein Element $x^\dagger \in X$ heißt

(i) *Ausgleichslösung* von $Tx = y$, wenn gilt

$$\|Tx^\dagger - y\|_Y = \min_{z \in X} \|Tz - y\|_Y;$$

(ii) *Minimum-Norm-Lösung* von $Tx = y$, wenn gilt

$$\|x^\dagger\|_X = \min \{\|z\|_X : z \text{ ist Ausgleichslösung von } Tx = y\}.$$

Offensichtlich ist für T bijektiv $x = T^{-1}y$ die einzige Ausgleichs- und damit auch Minimum-Norm-Lösung. Eine Ausgleichslösung muss aber nicht existieren, falls $\mathcal{R}(T)$ nicht abgeschlossen ist (da dann das Minimum in der Definition nicht angenommen wird). Um zu untersuchen, für welche $y \in Y$ eine Minimum-Norm-Lösung existiert, führen wir einen Operator ein, der y auf die Minimum-Norm-Lösung abbildet; diesen bezeichnen wir als *verallgemeinerte Inverse* oder *Pseudoinverse*. Dazu schränken wir zuerst Definitionsbereich und Bild von T ein, so dass der Operator invertierbar wird; dann erweitern wir die Inverse des eingeschränkten Operators auf ihren maximalen Definitionsbereich.

Satz 3.3. *Sei $T \in \mathcal{L}(X, Y)$ und setze*

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

Dann existiert eine eindeutige lineare Fortsetzung $T^\dagger$, genannt Moore–Penrose-Inverse, *von $\tilde{T}^{-1}$ mit*

$$\mathcal{D}(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp, \tag{3.1}$$

$$\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp. \tag{3.2}$$

Beweis. Wegen der Einschränkung auf $\mathcal{N}(T)^\perp$ und $\mathcal{R}(T)$ ist $\tilde{T}$ injektiv und surjektiv, daher existiert $\tilde{T}^{-1}$ und ist linear. Damit ist $T^\dagger$ auf $\mathcal{R}(T)$ wohldefiniert und linear. Für beliebige $y \in \mathcal{D}(T^\dagger)$ existieren aufgrund der Orthogonalzerlegung eindeutige $y_1 \in \mathcal{R}(T)$ und $y_2 \in \mathcal{R}(T)^\perp$ mit $y = y_1 + y_2$. Wegen $\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp$ ist also durch

$$T^\dagger y := T^\dagger y_1 + T^\dagger y_2 = T^\dagger y_1 = \tilde{T}^{-1} y_1 \tag{3.3}$$

eine eindeutige lineare Fortsetzung festgelegt. Damit ist $T^\dagger$ auf ganz $\mathcal{D}(T^\dagger)$ wohldefiniert. □

Ist T bijektiv, so gilt offensichtlich $T^\dagger = T^{-1}$. Beachten Sie, dass $T^\dagger$ nicht unbedingt eine *stetige* Fortsetzung sein muss!

Im weiteren Verlauf brauchen wir die folgenden Eigenschaften der Moore–Penrose-Inversen.

Lemma 3.4. *Die Moore–Penrose-Inverse $T^\dagger$ erfüllt $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$ sowie die* Moore–Penrose-Gleichungen

(i) $TT^\dagger T = T$,

(ii) $T^\dagger T T^\dagger = T^\dagger$,

(iii) $T^\dagger T = \mathrm{Id} - P_{\mathcal{N}}$,

(iv) $TT^\dagger = (P_{\overline{\mathcal{R}}})|_{\mathcal{D}(T^\dagger)}$,

wobei $P_{\mathcal{N}}$ und $P_{\overline{\mathcal{R}}}$ die orthogonalen Projektionen auf $\mathcal{N}(T)$ respektive $\overline{\mathcal{R}(T)}$ bezeichnen.

Beweis. Wir zeigen zuerst $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$. Nach Definition von $T^\dagger$ und (3.3) gilt für alle $y \in \mathcal{D}(T^\dagger)$

$$T^\dagger y = T^\dagger P_{\overline{\mathcal{R}}} y = \tilde{T}^{-1} P_{\overline{\mathcal{R}}} y, \tag{3.4}$$

denn für $y \in \mathcal{D}(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ ist $P_{\overline{\mathcal{R}}} y \in \mathcal{R}(T)$ (und nicht nur in $\overline{\mathcal{R}(T)}$ – diese fundamentale Eigenschaft werden wir in Folge öfter verwenden). Also ist $T^\dagger y \in \mathcal{R}(\tilde{T}^{-1}) = \mathcal{N}(T)^\perp$, d. h. $\mathcal{R}(T^\dagger) \subset \mathcal{N}(T)^\perp$. Umgekehrt gilt für alle $x \in \mathcal{N}(T)^\perp$, dass $T^\dagger T x = \tilde{T}^{-1}\tilde{T} x = x$ ist, d. h. $x \in \mathcal{R}(T^\dagger)$. Also ist $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$.

Zu (iv): Für $y \in \mathcal{D}(T^\dagger)$ folgt weiterhin aus (3.4) und $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$

$$TT^\dagger y = T\tilde{T}^{-1} P_{\overline{\mathcal{R}}} y = \tilde{T}\tilde{T}^{-1} P_{\overline{\mathcal{R}}} y = P_{\overline{\mathcal{R}}} y$$

wegen $\tilde{T}^{-1} P_{\overline{\mathcal{R}}} y \in \mathcal{N}(T)^\perp$ und $T = \tilde{T}$ auf $\mathcal{N}(T)^\perp$.

Zu (iii): Nach Definition von $T^\dagger$ gilt $T^\dagger T x = \tilde{T}^{-1} T x$ für alle $x \in X$ und deshalb

$$T^\dagger T x = \tilde{T}^{-1} T \left(P_{\mathcal{N}} x + (\mathrm{Id} - P_{\mathcal{N}}) x\right) = \tilde{T}^{-1} T P_{\mathcal{N}} x + \tilde{T}^{-1}\tilde{T}(\mathrm{Id} - P_{\mathcal{N}}) x = (\mathrm{Id} - P_{\mathcal{N}}) x.$$

Zu (ii): Einsetzen von (iv) und (3.4) ergibt für $y \in \mathcal{D}(T^\dagger)$

$$T^\dagger T T^\dagger y = T^\dagger P_{\overline{\mathcal{R}}} y = T^\dagger y.$$

Zu (i): Aus (iii) folgt sofort

$$TT^\dagger T x = T(\mathrm{Id} - P_{\mathcal{N}}) x = Tx - TP_{\mathcal{N}} x = Tx \qquad \text{für alle } x \in X. \qquad \square$$

(Tatsächlich sind die Moore–Penrose-Gleichungen eine eindeutige Charakterisierung von $T^\dagger$.)

Nun können wir zeigen, dass die Moore–Penrose-Inverse tatsächlich die Minimum-Norm-Lösung liefert; dabei charakterisieren wir auch gleich die Ausgleichslösungen.

Satz 3.5. *Sei $y \in \mathcal{D}(T^\dagger)$. Dann hat $Tx = y$*

(i) eine Ausgleichslösung, und die Ausgleichslösungen sind genau die Lösungen von

$$Tx = P_{\overline{\mathcal{R}}} y, \tag{3.5}$$

(ii) eine eindeutige Minimum-Norm-Lösung $x^\dagger \in X$, die gegeben ist durch

$$x^\dagger = T^\dagger y.$$

Die Menge aller Ausgleichslösungen ist dann gegeben durch $x^\dagger + \mathcal{N}(T)$.

Beweis. Zuerst existiert wegen $P_{\overline{\mathcal{R}}} y \in \mathcal{R}(T)$ für $y \in \mathcal{D}(T^\dagger)$ mindestens eine Lösung von (3.5). Aus der Optimalität der orthogonalen Projektion folgt weiter für jede solche Lösung $z \in X$, dass

$$\|Tz - y\|_Y = \|P_{\overline{\mathcal{R}}} y - y\|_Y = \min_{w \in \overline{\mathcal{R}(T)}} \|w - y\|_Y \leq \|Tx - y\|_Y \qquad \text{für alle } x \in X,$$

d. h. alle Lösungen von (3.5) sind Ausgleichslösungen von $Tx = y$. Umgekehrt gilt für jede Ausgleichslösung $z \in X$, dass

$$\|P_{\overline{\mathcal{R}}} y - y\|_Y \leq \|Tz - y\|_Y = \min_{x \in X} \|Tx - y\|_Y = \min_{w \in \mathcal{R}(T)} \|w - y\|_Y \leq \|P_{\overline{\mathcal{R}}} y - y\|_Y$$

wegen $P_{\overline{\mathcal{R}}} y \in \mathcal{R}(T)$ und damit $Tz = P_{\overline{\mathcal{R}}} y$. Damit ist (i) gezeigt.

Die Ausgleichslösungen sind also genau die Lösungen der linearen Gleichung $Tx = P_{\overline{\mathcal{R}}} y$, die sich eindeutig darstellen lassen als $x = \bar{x} + x_0$ mit $\bar{x} \in \mathcal{N}(T)^\perp$ und $x_0 \in \mathcal{N}(T)$. Da T injektiv auf $\mathcal{N}(T)^\perp$ ist, muss $\bar{x}$ unabhängig von x sein (sonst wäre $Tx' = T\bar{x}' \neq T\bar{x} = P_{\overline{\mathcal{R}}} y$ für $x' = \bar{x}' + x_0$ mit $\bar{x}' \neq \bar{x}$). Wegen

$$\|x\|_X^2 = \|\bar{x} + x_0\|_X^2 = \|\bar{x}\|_X^2 + 2\,(\bar{x}, x_0)_X + \|x_0\|_X^2 = \|\bar{x}\|_X^2 + \|x_0\|_X^2 \geq \|\bar{x}\|_X^2$$

ist $x^\dagger := \bar{x} \in \mathcal{N}(T)^\perp$ die eindeutige Minimum-Norm-Lösung.

Schließlich folgt aus $x^\dagger \in \mathcal{N}(T)^\perp$ und $Tx^\dagger = P_{\overline{\mathcal{R}}} y$ zusammen mit Lemma 3.4 (iii) und (ii)

$$x^\dagger = P_{\mathcal{N}^\perp} x^\dagger = (\mathrm{Id} - P_{\mathcal{N}}) x^\dagger = T^\dagger T x^\dagger = T^\dagger P_{\overline{\mathcal{R}}} y = T^\dagger T T^\dagger y = T^\dagger y,$$

was für (ii) zu zeigen war. □

Wir können eine alternative Charakterisierung angeben.

Folgerung 3.6. *Sei $y \in \mathcal{D}(T^\dagger)$. Dann ist $x \in X$ Ausgleichslösung von $Tx = y$ genau dann, wenn x die* Normalengleichungen

$$T^* T x = T^* y \tag{3.6}$$

erfüllt. Ist zusätzlich $x \in \mathcal{N}(T)^\perp$, so ist $x = x^\dagger$.

Beweis. Nach Satz 3.5 (i) ist $x \in X$ Ausgleichslösung genau dann, wenn $Tx = P_{\overline{\mathcal{R}}} y$ gilt. Nun ist $Tx = P_{\overline{\mathcal{R}}} y$ äquivalent zu $Tx \in \overline{\mathcal{R}(T)}$ und $Tx - y \in \overline{\mathcal{R}(T)}^{\perp} = \mathcal{N}(T^*)$, woraus $T^*(Tx - y) = 0$ folgt. Schließlich hat nach Satz 3.5 (ii) eine Ausgleichslösung x minimale Norm genau dann, wenn $x = T^\dagger y \in \mathcal{N}(T)^\perp$ gilt. □

Die Minimum-Norm-Lösung $x^\dagger$ von $Tx = y$ ist also auch die Lösung – und damit insbesondere Ausgleichslösung – von (3.6) mit minimaler Norm, d. h.

$$x^\dagger = (T^*T)^\dagger T^* y. \tag{3.7}$$

Zur Berechnung von $x^\dagger$ können wir also die Minimum-Norm-Lösung von (3.6) an Stelle von $Tx = y$ heranziehen, was manchmal vorteilhafter sein wird.

Bislang haben wir die Pseudo-Inverse auf ihrem Definitionsbereich betrachtet, ohne diesen näher zu untersuchen; dies holen wir nun nach. Nach Konstruktion ist $\mathcal{D}(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$. Da orthogonale Komplemente stets abgeschlossen sind, gilt

$$\overline{\mathcal{D}(T^\dagger)} = \overline{\mathcal{R}(T)} \oplus \mathcal{R}(T)^\perp = \mathcal{N}(T^*)^\perp \oplus \mathcal{N}(T^*) = Y,$$

d. h. $\mathcal{D}(T^\dagger)$ ist dicht in Y. Ist $\mathcal{R}(T)$ abgeschlossen, so ist daher $\mathcal{D}(T^\dagger) = Y$ (woraus umgekehrt folgt, dass $\mathcal{R}(T)$ abgeschlossen ist). Weiterhin ist für $y \in \mathcal{R}(T)^\perp = \mathcal{N}(T^\dagger)$ die Minimum-Norm-Lösung stets durch $x^\dagger = 0$ gegeben. Die zentrale Frage ist also, ob ein gegebenes $y \in \overline{\mathcal{R}(T)}$ auch in $\mathcal{R}(T)$ liegt. Gilt dies stets, so muss $T^\dagger$ sogar stetig sein. Umgekehrt reicht bereits die Existenz *eines* $y \in \overline{\mathcal{R}(T)} \setminus \mathcal{R}(T)$, dass $T^\dagger$ nicht stetig sein kann.

Satz 3.7. *Sei $T \in \mathcal{L}(X, Y)$. Dann ist $T^\dagger : \mathcal{D}(T^\dagger) \to X$ stetig genau dann, wenn $\mathcal{R}(T)$ abgeschlossen ist.*

Beweis. Wir wenden den Satz 1.4 vom abgeschlossenen Graphen an. Zuerst zeigen wir, dass $T^\dagger$ abgeschlossen ist. Sei $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T^\dagger)$ eine Folge mit $y_n \to y \in Y$ und $T^\dagger y_n \to x \in X$. Aus Lemma 3.4 (iv) folgt nun

$$TT^\dagger y_n = P_{\overline{\mathcal{R}}} y_n \to P_{\overline{\mathcal{R}}} y$$

wegen der Stetigkeit der orthogonalen Projektion. Zusammen mit der Stetigkeit von T folgt daraus

$$P_{\overline{\mathcal{R}}} y = \lim_{n\to\infty} P_{\overline{\mathcal{R}}} y_n = \lim_{n\to\infty} TT^\dagger y_n = Tx, \tag{3.8}$$

d. h. x ist eine Ausgleichslösung. Weiterhin gilt wegen $T^\dagger y_n \in \mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$ auch

$$T^\dagger y_n \to x \in \mathcal{N}(T)^\perp,$$

denn $\mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$ ist abgeschlossen. Nach Satz 3.5 (ii) bedeutet das, dass x die Minimum-Norm-Lösung von $Tx = y$ ist, d. h. $x = T^\dagger y$. Damit ist $T^\dagger$ abgeschlossen.

Ist nun $\mathcal{R}(T)$ abgeschlossen, so gilt $\mathcal{D}(T^\dagger) = Y$. Nach Satz 1.4 ist deshalb $T^\dagger : Y \to X$ stetig. Ist andererseits $T^\dagger$ stetig auf $\mathcal{D}(T^\dagger)$, so kann wegen der Dichtheit von $\mathcal{D}(T^\dagger)$ in Y der Operator $T^\dagger$ stetig fortgesetzt werden auf Y zu einem $\overline{T^\dagger} \in \mathcal{L}(Y, X)$ durch

$$\overline{T^\dagger} y := \lim_{n\to\infty} T^\dagger y_n \quad \text{für eine Folge } \{y_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T^\dagger) \text{ mit } y_n \to y \in Y.$$

(Wegen der Beschränktheit bildet $T^\dagger$ Cauchyfolgen auf Cauchyfolgen ab, weshalb $\overline{T^\dagger}$ wohldefiniert und stetig ist.) Sei nun $y \in \overline{\mathcal{R}(T)}$ und $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{R}(T)$ mit $y_n \to y$. Analog zu (3.8) folgt dann

$$y = P_{\overline{\mathcal{R}}} y = \lim_{n\to\infty} P_{\overline{\mathcal{R}}} y_n = \lim_{n\to\infty} TT^\dagger y_n = T\overline{T^\dagger} y \in \mathcal{R}(T),$$

und damit $\overline{\mathcal{R}(T)} = \mathcal{R}(T)$. □

Entsprechend nennt man die Operatorgleichung $Tx = y$ auch *schlecht gestellt nach Nashed*, wenn $\mathcal{R}(T)$ nicht abgeschlossen ist. Unglücklicherweise schließt das bereits den interessanten Fall von kompakten Operatoren aus.

Folgerung 3.8. *Sei $K \in \mathcal{K}(X, Y)$ mit unendlichdimensionalem Bild $\mathcal{R}(K)$. Dann ist $K^\dagger$ nicht stetig.*

Beweis. Angenommen, $K^\dagger$ ist stetig. Dann ist $\mathcal{R}(K)$ nach Satz 3.7 abgeschlossen, und damit hat der wie in Satz 3.3 definierte, nach Konstruktion bijektive, Operator $\tilde{K} := K : \mathcal{N}(K)^\perp \to \mathcal{R}(K)$ eine stetige Inverse $\tilde{K}^{-1} \in \mathcal{L}(\mathcal{R}(K), \mathcal{N}(K)^\perp)$. Nun ist K kompakt, und daher auch das Produkt $K \circ \tilde{K}^{-1}$. Wegen

$$K\tilde{K}^{-1} y = y \qquad \text{für alle } y \in \mathcal{R}(K)$$

ist dann aber auch die Identität $\mathrm{Id} : \mathcal{R}(K) \to \mathcal{R}(K)$ kompakt, und das ist nur möglich, wenn $\mathcal{R}(K)$ endlichdimensional ist. □

Für kompakte Operatoren müssen wir die Bedingung (iii) in der Definition der Korrektgestelltheit also mit anderen Mitteln erreichen. Dies werden wir im nächsten Kapitel tun.

3.2 SINGULÄRWERTZERLEGUNG KOMPAKTER OPERATOREN

Wir charakterisieren nun die Moore–Penrose-Inverse eines kompakten Operators $K \in \mathcal{K}(X, Y)$ mit Hilfe von Orthonormalsystemen. Dafür möchten wir eine Spektralzerlegung verwenden, die für nicht selbstadjungierte Operatoren aber nicht existiert; wegen Folgerung 3.6 können wir stattdessen aber genauso gut K^*K betrachten. Dies führt auf die *Singulärwertzerlegung*.

Satz 3.9. *Sei $K \in \mathcal{K}(X, Y)$. Dann existieren*

(i) eine Nullfolge $\{\sigma_n\}_{n\in\mathbb{N}}$ mit $\sigma_1 \geq \sigma_2 \geq \cdots > 0$,

(ii) eine Orthonormalbasis $\{u_n\}_{n\in\mathbb{N}} \subset Y$ von $\overline{\mathcal{R}(K)}$,

(iii) eine Orthonormalbasis $\{v_n\}_{n\in\mathbb{N}} \subset X$ von $\overline{\mathcal{R}(K^*)}$

(die eventuell endlich sind) mit

$$Kv_n = \sigma_n u_n \quad \text{und} \quad K^*u_n = \sigma_n v_n \qquad \text{für alle } n \in \mathbb{N} \tag{3.9}$$

und

$$Kx = \sum_{n\in\mathbb{N}} \sigma_n \,(x, v_n)_X \, u_n \qquad \text{für alle } x \in X. \tag{3.10}$$

Eine Folge $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$, für die die Singulärwertzerlegung (3.10) *gilt, heißt* singuläres System.

Beweis. Da $K^*K : X \to X$ kompakt und selbstadjungiert ist, liefert der Spektralsatz (Satz 2.2) eine Nullfolge $\{\lambda_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ (nach absteigendem Betrag sortiert) und ein Orthonormalsystem $\{v_n\}_{n\in\mathbb{N}} \subset X \setminus \{0\}$ aus zugehörigen Eigenvektoren mit

$$K^*Kx = \sum_{n\in\mathbb{N}} \lambda_n \,(x, v_n)_X \, v_n \qquad \text{für alle } x \in X.$$

Wegen $\lambda_n = \lambda_n \|v_n\|_X^2 = (\lambda_n v_n, v_n)_X = (K^*Kv_n, v_n)_X = \|Kv_n\|_Y^2 > 0$ können wir für alle $n \in \mathbb{N}$

$$\sigma_n := \sqrt{\lambda_n} > 0 \qquad \text{und} \qquad u_n := \sigma_n^{-1} K v_n \in Y$$

definieren. (Bricht die Folge $\{v_n\}_{n\in\mathbb{N}}$ ab, so setzen wir $\sigma_n := 0$ für $v_n = 0$.) Letztere bilden aufgrund von

$$\big(u_i, u_j\big)_Y = \frac{1}{\sigma_i\sigma_j}\big(Kv_i, Kv_j\big)_Y = \frac{1}{\sigma_i\sigma_j}\big(K^*Kv_i, v_j\big)_X = \frac{\lambda_i}{\sigma_i\sigma_j}\big(v_i, v_j\big)_X = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst,} \end{cases}$$

ein Orthonormalsystem. Weiterhin gilt für alle $n \in \mathbb{N}$

$$K^*u_n = \sigma_n^{-1} K^*Kv_n = \sigma_n^{-1}\lambda_n v_n = \sigma_n v_n.$$

Nun ist $\{v_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K^*K)}$. Weiterhin gilt $\overline{\mathcal{R}(K^*K)} = \overline{\mathcal{R}(K^*)}$, denn für $x \in \overline{\mathcal{R}(K^*)}$ existiert eine Folge $\{y_n\}_{n\in\mathbb{N}} \subset Y$ mit $K^* y_n \to x$; insbesondere können wir $y_n \in \mathcal{N}(K^*)^\perp = \overline{\mathcal{R}(K)}$ annehmen, und ein Diagonalfolgenargument ergibt $x \in \overline{\mathcal{R}(K^*K)}$. (Die andere Richtung ist klar.) Also ist $\{v_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$, und damit gilt

$$Kx = KP_{\mathcal{N}^\perp} x = K\left(\sum_{n\in\mathbb{N}} (x, v_n)_X \, v_n\right) \qquad \text{für alle } x \in X.$$

Die Singulärwertzerlegung (3.10) erhalten wir nun daraus, indem wir K durch die entsprechende Basisdarstellung „durchreichen". Da wir das in Folge öfter anwenden werden, begründen wir ausführlich, dass wir das dürfen. Betrachte dafür $x_N := \sum_{n=1}^{N} (x, v_n)_X \, v_n$ für $x \in X$ und $N \in \mathbb{N}$ beliebig. Dann gilt offensichtlich $x_N \to P_{\mathcal{N}^\perp} x$ für $N \to \infty$ und damit wegen der Stetigkeit von K auch

$$\begin{aligned} Kx = K\,(P_{\mathcal{N}^\perp} x) = K(\lim_{N\to\infty} x_N) &= \lim_{N\to\infty} Kx_N \\ = \lim_{N\to\infty} \sum_{n=1}^{N} (x, v_n)_X \, Kv_n &= \sum_{n\in\mathbb{N}} (x, v_n)_X \, Kv_n \end{aligned} \tag{3.11}$$

nach Definition der Reihe. Also gilt für alle $x \in X$

$$Kx = \sum_{n\in\mathbb{N}} (x, v_n)_X \, Kv_n = \sum_{n\in\mathbb{N}} (x, v_n)_X \, \sigma_n u_n = \sum_{n\in\mathbb{N}} (x, K^* u_n)_X \, u_n = \sum_{n\in\mathbb{N}} (Kx, u_n)_X \, u_n.$$

Aus der zweiten Gleichung folgt (3.10); die letzte impliziert, dass $\{u_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K)}$ ist. □

Da Eigenwerte von K^*K mit Eigenvektor v_n auch Eigenwerte von KK^* mit Eigenvektor u_n sind, erhält man mit Hilfe von (3.9) auch ein singuläres System $\{(\sigma_n, v_n, u_n)\}_{n\in\mathbb{N}}$ von K^* mit

$$K^* y = \sum_{n\in\mathbb{N}} \sigma_n \, (y, u_n)_Y \, v_n \qquad \text{für alle } y \in Y. \tag{3.12}$$

Außerdem ist natürlich $\{(\sigma_n^2, v_n, v_n)\}_{n\in\mathbb{N}}$ nach Konstruktion ein singuläres System von K^*K.

Wir verwenden nun die Singulärwertzerlegung von K, um den Definitionsbereich $\mathcal{D}(K^\dagger) = \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ der Moore–Penrose-Inversen $K^\dagger$ zu charakterisieren. Wie bereits vor Satz 3.7 diskutiert, reduziert sich dies auf die Frage, wann $y \in \overline{\mathcal{R}(K)}$ tatsächlich in $\mathcal{R}(K)$ liegt.

Satz 3.10. *Sei $K \in \mathcal{K}(X, Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ und $y \in \overline{\mathcal{R}(K)}$. Dann ist $y \in \mathcal{R}(K)$ genau dann, wenn die* Picard-Bedingung

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2} |\, (y, u_n)_Y \,|^2 < \infty \tag{3.13}$$

erfüllt ist.

In diesem Fall gilt

$$K^\dagger y = \sum_{n\in\mathbb{N}} \sigma_n^{-1} (y, u_n)_Y \, v_n. \tag{3.14}$$

Beweis. Sei $y \in \mathcal{R}(K)$, es existiere also ein $x \in X$ mit $Kx = y$. Dann ist

$$(y, u_n)_Y = (x, K^* u_n)_X = \sigma_n (x, v_n)_X \qquad \text{für alle } n \in \mathbb{N},$$

und damit folgt aus der Besselschen Ungleichung (2.2)

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2} |(y, u_n)_Y|^2 = \sum_{n\in\mathbb{N}} |(x, v_n)_X|^2 \le \|x\|_X^2 < \infty.$$

Sei umgekehrt $y \in \overline{\mathcal{R}(K)}$ und gelte (3.13). Insbesondere ist damit $\{\sum_{n=1}^N \sigma_n^{-2} |(y, u_n)_Y|^2\}_{N\in\mathbb{N}}$ eine Cauchy-Folge. Dann ist auch $\{x_N\}_{N\in\mathbb{N}}$ mit

$$x_N := \sum_{n=1}^N \sigma_n^{-1} (y, u_n)_Y \, v_n$$

eine Cauchyfolge, denn die Folge $\{v_n\}_{n\in\mathbb{N}}$ bildet ein Orthonormalsystem und daher gilt

$$\|x_N - x_M\|_X^2 = \|\textstyle\sum_{n=N+1}^M \sigma_n^{-1} (y, u_n)_Y \, v_n\|_X^2 = \sum_{n=N+1}^M |\sigma_n^{-1} (y, u_n)_Y|^2 \to 0 \quad \text{für } N, M \to \infty.$$

Weiterhin ist $\{v_n\}_{n\in\mathbb{N}} \subset \overline{\mathcal{R}(K^*)}$. Also konvergiert $\{x_N\}_{N\in\mathbb{N}} \subset \overline{\mathcal{R}(K^*)}$ gegen ein

$$x := \sum_{n\in\mathbb{N}} \sigma_n^{-1} (y, u_n)_Y \, v_n \in X,$$

für das wegen der Abgeschlossenheit von $\overline{\mathcal{R}(K^*)}$ auch gilt $x \in \overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$. Nun gilt wie in (3.11)

$$Kx = \sum_{n\in\mathbb{N}} \sigma_n^{-1} (y, u_n)_Y \, K v_n = \sum_{n\in\mathbb{N}} (y, u_n)_Y \, u_n = P_{\overline{\mathcal{R}}} y = y,$$

woraus $y \in \mathcal{R}(K)$ folgt. Nach Satz 3.5 ist aber $Kx = P_{\overline{\mathcal{R}}} y$ und $x \in \mathcal{N}(K)^\perp$ äquivalent zu $x = K^\dagger y$. □

Die Picard-Bedingung sagt, dass eine Minimum-Norm-Lösung nur existieren kann, wenn die „Fourier-Koeffizienten“ $(y, u_n)_Y$ von y im Vergleich zu den Singulärwerten σ_n schnell genug abfallen. Die Darstellung (3.14) zeigt auch, wie Störungen in y sich auf Störungen in $x^\dagger$ auswirken: Ist $y^\delta = y + \delta u_n$ für ein $\delta > 0$ und $n \in \mathbb{N}$, so gilt

$$\|K^\dagger y^\delta - K^\dagger y\|_X = \delta \|K^\dagger u_n\|_X = \sigma_n^{-1} \delta \to \infty \quad \text{für } n \to \infty,$$

und je schneller die Singulärwerte abfallen, desto mehr wird der Fehler für festes n verstärkt. Man unterscheidet daher

- *moderat schlecht gestellte* Probleme, für die $c, r > 0$ existieren so das $\sigma_n \geq cn^{-r}$ für alle $n \in \mathbb{N}$ gilt (d. h. σ_n höchstens polynomiell gegen Null geht), und
- *stark schlecht gestellte* Probleme, für die dies nicht der Fall ist. Gilt sogar $\sigma_n \leq ce^{-n^r}$ für alle $n \in \mathbb{N}$ und $c, r > 0$ (d. h. σ_n geht mindestens exponentiell gegen Null), so ist das Problem *exponentiell schlecht gestellt.*

Für exponentiell schlecht gestellte Probleme kann man in der Regel keine Lösung erwarten, die über eine sehr grobe Näherung hinausgeht. Ist aber $\mathcal{R}(K)$ endlichdimensional, so bricht die Folge $\{\sigma_n\}_{n\in\mathbb{N}}$ ab und der Fehler bleibt beschränkt; in diesem Fall ist $K^\dagger$ also wie erwartet stetig.

Die Singulärwertzerlegung ist ein wertvolles analytisches Werkzeug; die explizite Bestimmung für konkrete Operatoren ist in der Regel aber sehr aufwendig. Wir betrachten wieder die Differentiation als einfaches Beispiel.

Beispiel 3.11. Sei $X = L^2(\Omega)$ für $\Omega = (0,1)$ und $K \in \mathcal{K}(X,X)$ ein Integraloperator gegeben durch

$$[Kx](t) = \int_0^1 k(s,t)x(s)\,ds \qquad \text{mit} \qquad k(s,t) = \begin{cases} 1 & \text{falls } s \leq t, \\ 0 & \text{sonst.} \end{cases}$$

Ist $x = y'$ für ein $y \in C^1([0,1])$ mit $y(0) = 0$, so gilt dann

$$[Kx](t) = \int_0^t x(s)\,ds = y(t) - y(0) = y(t),$$

d. h. für $y \in C^1([0,1])$ ist die Ableitung y' eine Lösung der Operatorgleichung $Kx = y$ (die zwar auch für $y \in L^2(\Omega)$ sinnvoll formuliert ist, dort aber keine Lösung haben muss).

Der zugehörige adjungierte Operator ist dann gegeben durch

$$[K^*y](t) = \int_0^1 k(t,s)y(s)\,ds = \int_t^1 y(s)\,ds,$$

denn $k(t,s) = 1$ für $s \geq t$ und 0 sonst. Wir suchen nun die Eigenwerte und Eigenvektoren von K^*K, d. h. $\lambda > 0$ und $v \in L^2(\Omega)$ mit

$$\lambda v(t) = [K^*Kv](t) = \int_t^1 \int_0^s v(r)\,dr\,ds. \tag{3.15}$$

Wir gehen zunächst formal vor. Einsetzen von $t = 1$ liefert $\lambda v(1) = 0$ und daher $v(1) = 0$. Ableiten von (3.15) ergibt

$$\lambda v'(t) = \frac{d}{dt}\left(-\int_1^t \int_0^s v(r)\,dr\,ds\right) = -\int_0^t v(r)\,dr,$$

woraus durch Einsetzen von $t = 0$ folgt, dass $v'(0) = 0$ sein muss. Ein weiteres Mal Ableiten führt nun auf die gewöhnliche Differentialgleichung

$$\lambda v''(t) + v(t) = 0,$$

welche die allgemeine Lösung

$$v(t) = c_1 \sin(\sigma^{-1} t) + c_2 \cos(\sigma^{-1} t)$$

für $\sigma := \sqrt{\lambda}$ und noch zu bestimmende Konstanten c_1, c_2 hat. Einsetzen in die Randbedingungen $v'(0) = 0$ und $v(1) = 0$ führt auf $c_1 = 0$ beziehungsweise $c_2 \cos(\sigma^{-1}) = 0$. Da $c_2 = 0$ auf die triviale Lösung $v = 0$ führen würde (und der Nullvektor ja kein Eigenvektor ist), muss $\cos(\sigma^{-1}) = 0$ sein; für die Singulärwerte σ_n kommen also nur die Kehrwerte der Nullstellen des Kosinus in Frage, nämlich

$$\sigma_n = \frac{2}{(2n-1)\pi}, \qquad n \in \mathbb{N}.$$

Damit sind die Eigenvektoren

$$v_n(t) = \sqrt{2} \cos\left((n - \tfrac{1}{2})\pi\, t\right), \qquad n \in \mathbb{N},$$

wobei wir die Konstante $c_2 = \sqrt{2}$ gewählt haben, um $\|v_n\|_{L^2(\Omega)} = 1$ zu erreichen. Für u_n berechnen wir

$$u_n = \sigma_n^{-1} K v_n = (n - \tfrac{1}{2})\pi \int_0^t \sqrt{2} \cos\left((n - \tfrac{1}{2})\pi\, s\right) ds = \sqrt{2} \sin\left((n - \tfrac{1}{2})\pi\, t\right), \quad n \in \mathbb{N}.$$

Nun gilt $v_n, u_n \in L^2(\Omega)$, und man verifiziert leicht, dass σ_n^2 und v_n die Eigenwertgleichung (3.15) erfüllen. Wie im Beweis von Satz 3.9 erhält man daraus die Singulärwertzerlegung von K, so dass $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System bildet.

Wegen $\sigma_n = O(\frac{1}{n})$ ist Ableiten (in dieser Formulierung) also ein moderat schlecht gestelltes Problem. Man kann nun zeigen, dass $\{u_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $L^2(\Omega)$ ist (die ja nicht eindeutig sein muss). Die Picard-Bedingung (3.13) für ein $y \in L^2(\Omega)$ ist dann gegeben durch

$$\sum_{n\in\mathbb{N}} \frac{1}{4}(2n-1)^2\pi^2 |\,(y, u_n)_{L^2}\,|^2 < \infty.$$

Weiterhin folgt aus (2.3) auch

$$y = \sum_{n\in\mathbb{N}} (y, u_n)_{L^2}\, u_n,$$

und gliedweises Differenzieren der Fourier-Reihe auf der rechten Seite ergibt dann (formal)

$$z := \sum_{n\in\mathbb{N}} (y, u_n)_{L^2}\, u_n' = \sum_{n\in\mathbb{N}} \left(n - \tfrac{1}{2}\right) \pi\, (y, u_n)_{L^2}\, v_n.$$

Die Picard-Bedingung ist also äquivalent zu der Bedingung, dass $\|z\|^2_{L^2(\Omega)} < \infty$ ist und damit die gliedweise differenzierte Reihe (in $L^2(\Omega)$) konvergiert; in diesem Fall ist $K^\dagger y = z$. Für stetig differenzierbare y ist die Konvergenz sogar gleichmäßig, und es gilt $y' = z = K^\dagger y$.

Mit Hilfe der Singulärwertzerlegung lassen sich auch Funktionen von kompakten Operatoren definieren, was später hilfreich sein wird. Sei $\varphi : [0, \infty) \to \mathbb{R}$ eine stückweise stetige und lokal beschränkte Funktion. Dann definieren wir für $K \in \mathcal{K}(X, Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ den Operator $\varphi(K^*K) : X \to X$ durch

$$\varphi(K^*K)x = \sum_{n\in\mathbb{N}} \varphi(\sigma_n^2)\,(x, v_n)_X\, v_n + \varphi(0)P_{\mathcal{N}(K)}x \qquad \text{für alle } x \in X. \tag{3.16}$$

Diese Reihe konvergiert in X, denn φ wird nur auf dem abgeschlossenen und beschränkten Intervall $[0, \sigma_1^2] = [0, \|K\|^2_{\mathcal{L}(X,Y)}]$ ausgewertet. Weiterhin folgt aus der Besselschen Ungleichung

$$\|\varphi(K^*K)\|_{\mathcal{L}(X,X)} \le \sup_{n\in\mathbb{N}} |\varphi(\sigma_n^2)| + \varphi(0) \le 2 \sup_{\lambda\in[0,\|K\|^2_{\mathcal{L}(X,Y)}]} |\varphi(\lambda)| < \infty, \tag{3.17}$$

d. h. $\varphi(K^*K) \in \mathcal{L}(X, X)$.

Wir betrachten hier speziell Potenzfunktionen $\varphi(t) = t^r$ für $r \ge 0$ und insbesondere die folgenden Beispiele.

Beispiel 3.12. Sei $K \in \mathcal{K}(X, Y)$.

(i) Für $\varphi(t) = 1$ gilt $\varphi(K^*K) = \mathrm{Id}$, denn für alle $x \in X$ ist

$$\varphi(K^*K)x = \sum_{n\in\mathbb{N}} (x, v_n)_X\, v_n + P_{\mathcal{N}(K)}x = P_{\overline{\mathcal{R}(K^*)}}x + P_{\mathcal{N}(K)}x = x$$

wegen $\overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$.

(ii) Für $\varphi(t) = t$ ist wegen $\varphi(0) = 0$ und dem Spektralsatz $\varphi(K^*K) = K^*K$.

(iii) Für $\varphi(t) = \sqrt{t}$ nennen wir $|K| := \varphi(K^*K)$ den *Betrag* von K; es gilt also wegen $\sigma_n > 0$

$$|K|x = \sum_{n\in\mathbb{N}} \sigma_n\,(x, v_n)_X\, v_n \qquad \text{für alle } x \in X.$$

Vergleicht man das letzte Beispiel mit der Singulärwertzerlegung (3.10), so sieht man, dass sich $|K|$ im wesentlichen genauso verhält wie K, mit dem einzigen Unterschied, dass er nach X statt nach Y abbildet. Die folgenden Eigenschaften, die wir später benötigen werden, illustrieren dies.

Lemma 3.13. *Sei $K \in \mathcal{K}(X, Y)$. Dann gilt*

(i) $|K|^{r+s} = |K|^r \circ |K|^s$ *für alle* $r, s \ge 0$;

(ii) $|K|^r$ *ist selbstadjungiert für alle* $r \ge 0$;

(iii) $\||K|x\|_X = \|Kx\|_Y$ *für alle* $x \in X$;

(iv) $\mathcal{R}(|K|) = \mathcal{R}(K^*)$.

Beweis. Aussage (i) folgt direkt aus

$$\begin{aligned}|K|^{r+s}x &= \sum_{n\in\mathbb{N}} \sigma_n^{r+s}\,(x,v_n)_X\,v_n = \sum_{n\in\mathbb{N}} \sigma_n^r\left(\sigma_n^s\,(x,v_n)_X\right)v_n\\ &= \sum_{n\in\mathbb{N}} \sigma_n^r\left(\sum_{m\in\mathbb{N}} \sigma_m^s\,(x,v_m)_X\,v_m, v_n\right)_X v_n\\ &= \sum_{n\in\mathbb{N}} \sigma_n^r\,(|K|^s x, v_n)_X\,v_n = |K|^r(|K|^s x),\end{aligned}$$

da $\{v_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem ist.

Für beliebige $x, z \in X$ und $r \geq 0$ folgt weiterhin aus der Bilinearität und Symmetrie des Skalarprodukts

$$(|K|^r x, z)_X = \sum_{n\in\mathbb{N}} \sigma_n^r\,(x,v_n)_X\,(v_n,z)_X = (x,|K|^r z)_X$$

und damit (ii).

Aussage (iii) folgt aus (i), (ii), und

$$\||K|x\|_X^2 = (|K|x,|K|x)_X = \left(|K|^2 x, x\right)_X = (K^*Kx,x)_X = (Kx,Kx)_X = \|Kx\|_X^2.$$

Zu (iv): Sei $\{(\sigma_n,u_n,v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K. Dann ist $\{(\sigma_n,v_n,u_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K^* und – nach Definition – $\{(\sigma_n,v_n,v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von $|K|$. Nun ist $x \in \mathcal{R}(K^*)$ genau dann, wenn $Kx \in \mathcal{R}(KK^*)$ und $x \in \mathcal{N}(K)^\perp$ ist. Die Picard-Bedingung für $Kx \in \mathcal{R}(KK^*)$ ist

$$\infty > \sum_{n\in\mathbb{N}} \sigma_n^{-4}|\,(Kx,u_n)_Y\,|^2 = \sum_{n\in\mathbb{N}} \sigma_n^{-4}|\,(x,K^*u_n)_X\,|^2 = \sum_{n\in\mathbb{N}} \sigma_n^{-2}|\,(x,v_n)_X\,|^2.$$

Dies ist aber (vergleiche den Beweis von Satz 3.10) die Picard-Bedingung für $x \in \mathcal{R}(|K|)$, und für $x \in \mathcal{N}(K)^\perp$ ist diese Bedingung auch notwendig. □

Der Beweis der letzten Eigenschaft deutet bereits an, dass wir $|K|$ verwenden können, um eine Art Picard-Bedingung für $x \in \overline{\mathcal{R}(K^*)}$ (statt $y \in \overline{\mathcal{R}(K)}$) zu formulieren; diese werden wir in einem späteren Kapitel verwenden, um besonders gut approximierbare Minimum-Norm-Lösungen zu charakterisieren.

Schließlich benötigen wir noch die folgende Ungleichung.

Lemma 3.14. *Sei $K \in \mathcal{K}(X,Y)$. Dann gilt für alle $r > s \geq 0$ und $x \in X$ die* Interpolationsungleichung

$$\||K|^s x\|_X \leq \||K|^r x\|_X^{\frac{s}{r}}\,\|x\|_X^{1-\frac{s}{r}}. \tag{3.18}$$

Beweis. Nach Definition von $|K|^s$ gilt

$$\||K|^s x\|_X^2 = \sum_{n\in\mathbb{N}} \sigma_n^{2s} |(x, v_n)_X|^2, \tag{3.19}$$

woraus mit der Besselschen Ungleichung auch sofort die Aussage für $s = 0$ folgt.

Für $s > 0$ wenden wir die Höldersche Ungleichung

$$\sum_{n\in\mathbb{N}} a_n b_n \le \left(\sum_{n\in\mathbb{N}} a_n^p\right)^{\frac{1}{p}} \left(\sum_{n\in\mathbb{N}} b_n^q\right)^{\frac{1}{q}} \qquad \text{für} \quad \frac{1}{p} + \frac{1}{q} = 1$$

an auf

$$a_n := \sigma_n^{2s} |\,(x, v_n)_X\,|^{2\frac{s}{r}}, \qquad b_n := |\,(x, v_n)_X\,|^{2-2\frac{s}{r}}, \qquad p = \frac{r}{s}, \qquad q = \frac{r}{r-s}.$$

Dann folgt wieder aus (3.19) und der Besselschen Ungleichung

$$\begin{aligned} \||K|^s x\|_X^2 &\le \left(\sum_{n\in\mathbb{N}} \sigma_n^{2r} |\,(x, v_n)_X\,|^2\right)^{\frac{s}{r}} \left(\sum_{n\in\mathbb{N}} |\,(x, v_n)_X\,|^2\right)^{1-\frac{s}{r}} \\ &\le \||K|^r x\|_X^{2\frac{s}{r}} \, \|x\|_X^{2(1-\frac{s}{r})}, \end{aligned}$$

und Wurzelziehen liefert die Aussage. □

4 REGULARISIERUNGSVERFAHREN

Wie im letzten Kapitel gezeigt, existiert für $y \in \mathcal{D}(T^\dagger)$ eine eindeutige Minimum-Norm-Lösung $x^\dagger = T^\dagger y$ der schlechtgestellten Operatorgleichung $Tx = y$. In der Praxis hat man allerdings selten die "exakten Daten" y zur Hand, sondern nur eine "gestörte Messung" $y^\delta \in B_\delta(y)$, d. h. mit

$$\|y - y^\delta\|_Y \leq \delta,$$

wobei $\delta > 0$ das *Fehlerniveau* bezeichnet. Da $T^\dagger$ im Allgemeinen nicht stetig ist, liefert $T^\dagger y^\delta$ in der Regel keine gute Näherung für $x^\dagger$, selbst wenn zufällig $y^\delta \in \mathcal{D}(T^\dagger)$ gilt. Wir suchen daher eine Näherung x_α^δ, die einerseits stetig von y^δ – und damit von δ – abhängt, und andererseits durch Wahl des *Regularisierungsparameters* $\alpha > 0$ so nahe an $x^\dagger$ gebracht werden kann, wie das Fehlerniveau δ zulässt. Insbesondere soll für $\delta \to 0$ und geeignete Wahl von $\alpha(\delta)$ auch $x_{\alpha(\delta)}^\delta \to x^\dagger$ gelten. Ein Verfahren, das eine solche Näherung konstruiert, wird *Regularisierungsverfahren* genannt.

4.1 REGULARISIERUNG UND PARAMETERWAHL

Im Falle von linearen Operatoren im Hilbertraum lassen sich diese Konstruktionsverfahren in Form von *Regularisierungsoperatoren* definieren, die man als stetigen Ersatz für die unbeschränkte Pseudoinverse $T^\dagger$ auffassen kann. Dies führt auf die folgende Definition.

Definition 4.1. Sei $T \in \mathcal{L}(X, Y)$ ein beschränkter Operator zwischen den Hilberträumen X und Y. Eine Familie $\{R_\alpha\}_{\alpha>0}$ von linearen Operatoren $R_\alpha : Y \to X$ heißt *Regularisierung* (von $T^\dagger$), falls gilt

(i) $R_\alpha \in \mathcal{L}(Y, X)$ für alle $\alpha > 0$;

(ii) $R_\alpha y \to T^\dagger y$ für $\alpha \to 0$ und alle $y \in \mathcal{D}(T^\dagger)$.

Eine Regularisierung ist also eine punktweise Approximation der Moore–Penrose-Inversen durch stetige Operatoren. Aus dem Satz von Banach–Steinhaus folgt jedoch, dass die Konvergenz im Allgemeinen nicht gleichmäßig sein kann, wenn $T^\dagger$ nicht stetig ist.

Satz 4.2. *Sei $T \in \mathcal{L}(X,Y)$ und sei $\{R_\alpha\}_{\alpha>0} \subset \mathcal{L}(Y,X)$ eine Regularisierung. Ist $T^\dagger$ nicht stetig, so ist $\{R_\alpha\}_{\alpha>0}$ nicht gleichmäßig beschränkt. Insbesondere existiert dann ein $y \in Y$ und eine Nullfolge $\{\alpha_n\}_{n\in\mathbb{N}}$ mit $\|R_{\alpha_n} y\|_X \to \infty$.*

Beweis. Angenommen, es existiert kein solches $y \in Y$. Dann ist die Familie $\{R_\alpha\}_{\alpha>0} \subset \mathcal{L}(Y,X)$ punktweise und damit nach dem Satz 1.5 von Banach–Steinhaus gleichmäßig beschränkt. Es existiert daher ein $M > 0$ mit $\|R_\alpha\|_{\mathcal{L}(Y,X)} \leq M$ für alle $\alpha > 0$. Zusammen mit der punktweisen Konvergenz $R_\alpha \to T^\dagger$ auf der dichten Teilmenge $\mathcal{D}(T^\dagger) \subset Y$ ergibt dann Folgerung 1.6, dass diese Konvergenz sogar auf ganz $\overline{\mathcal{D}(T^\dagger)} = Y$ gilt. Folgerung 1.7 liefert dann, dass $T^\dagger$ stetig ist, und die Aussage folgt durch Kontraposition. □

Tatsächlich kann man unter einer Zusatzannahme zeigen, dass die Divergenz für *alle* $y \notin \mathcal{D}(T^\dagger)$ gelten muss.

Satz 4.3. *Sei $T \in \mathcal{L}(X,Y)$ mit $T^\dagger$ nicht stetig und sei $\{R_\alpha\}_{\alpha>0} \subset \mathcal{L}(Y,X)$ eine Regularisierung von $T^\dagger$. Ist*

$$\sup_{\alpha>0} \|TR_\alpha\|_{\mathcal{L}(Y,Y)} < \infty, \tag{4.1}$$

so gilt $\|R_\alpha y\|_X \to \infty$ für $\alpha \to 0$ und alle $y \notin \mathcal{D}(T^\dagger)$.

Beweis. Sei $y \in Y \setminus \mathcal{D}(T^\dagger) = \overline{\mathcal{R}(T)} \setminus \mathcal{R}(T)$ beliebig. Angenommen, es gibt eine Nullfolge $\{\alpha_n\}_{n\in\mathbb{N}}$ für die $\{R_{\alpha_n} y\}_{n\in\mathbb{N}}$ beschränkt ist. Dann existiert eine Teilfolge $\{x_k\}_{k\in\mathbb{N}}$, $x_k := R_{\alpha_{n_k}} y$, mit $x_k \rightharpoonup x \in X$. Da beschränkte lineare Operatoren stets schwach stetig sind, gilt dann auch $Tx_k \rightharpoonup Tx$.

Andererseits folgt aus der Stetigkeit von T und der punktweisen Konvergenz $R_\alpha \to T^\dagger$ auf $\mathcal{D}(T^\dagger)$ mit Lemma 3.4 (iv), dass $TR_\alpha \tilde{y} \to TT^\dagger \tilde{y} = P_{\overline{\mathcal{R}}} \tilde{y}$ für alle $\tilde{y} \in \mathcal{D}(T^\dagger)$ gilt. Die Annahme (4.1) und Folgerung 1.6 liefert dann die punktweise Konvergenz von $TR_{\alpha_n} \to P_{\overline{\mathcal{R}}}$ auf ganz Y. Aus $Tx_k = TR_{\alpha_{n_k}} y \to P_{\overline{\mathcal{R}}} y$ und $Tx_k \rightharpoonup Tx$ folgt nun mit der Eindeutigkeit des Grenzwertes $Tx = P_{\overline{\mathcal{R}}} y$. Also ist $P_{\overline{\mathcal{R}}} y \in \mathcal{R}(T)$ und damit $y \in \mathcal{R}(T)$, im Widerspruch zur Voraussetzung $y \notin \mathcal{R}(T)$. □

Nun ist in der Regel das gegebene $y^\delta \in B_\delta(y)$ nicht in $\mathcal{D}(T^\dagger)$. Wir betrachten daher den *Gesamtfehler*, den wir aufsplitten können in

$$\begin{aligned} \|R_\alpha y^\delta - T^\dagger y\|_X &\leq \|R_\alpha y^\delta - R_\alpha y\|_X + \|R_\alpha y - T^\dagger y\|_X \\ &\leq \delta \|R_\alpha\|_{\mathcal{L}(Y,X)} + \|R_\alpha y - T^\dagger y\|_X. \end{aligned} \tag{4.2}$$

Diese Zerlegung ist ein fundamentales Hilfsmittel in der Regularisierungstheorie, das uns noch öfter begegnen wird. Der erste Term beschreibt dabei den *(fortgepflanzten) Datenfehler*, der aufgrund von Satz 4.2 für $\alpha \to 0$ nicht beschränkt bleiben kann, solange

$\delta > 0$ ist. Der zweite Term beschreibt den *Verfahrensfehler*, der aufgrund der punktweisen Konvergenz für $\alpha \to 0$ gegen Null geht. Um eine sinnvolle Näherung zu erhalten, muss daher α in Abhängigkeit von δ korrekt gewählt werden. Insbesondere fordern wir, dass der Gesamtfehler für $\delta \to 0$ gegen Null geht. Hier und in Folge schreiben wir kurz $\mathbb{R}^+ := (0, \infty)$.

Definition 4.4. Eine Funktion $\alpha : \mathbb{R}^+ \times Y \to \mathbb{R}^+$, $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$, heißt *Parameterwahlstrategie*. Man unterscheidet

(i) *a priori-Strategien*, die nur von δ abhängen;

(ii) *a posteriori-Strategien*, die von δ und y^δ abhängen;

(iii) *heuristische Strategien*, die nur von y^δ abhängen.

Ist $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung von $T^\dagger$ und α eine Parameterwahlstrategie, so heißt das Paar (R_α, α) ein *(konvergentes) Regularisierungsverfahren*, falls gilt

$$\lim_{\delta \to 0} \sup_{y^\delta \in B_\delta(y)} \|R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y\|_X = 0 \qquad \text{für alle } y \in \mathcal{D}(T^\dagger). \tag{4.3}$$

Wir fordern also, dass für *alle* mit dem Fehlerniveau verträglichen Störungen y^δ der Gesamtfehler für verschwindendes Fehlerniveau gegen Null geht.

Wir betrachten zunächst die klassischen Beispiele für Parameterwahlstrategien.

A PRIORI-STRATEGIEN

Als erstes zeigen wir, dass für jede Regularisierung stets eine a priori-Strategie – und damit ein Regularisierungsverfahren – existiert.

Satz 4.5. *Sei $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung von $T^\dagger$. Dann existiert eine a priori-Strategie α, so dass (R_α, α) ein Regularisierungsverfahren ist.*

Beweis. Sei $y \in \mathcal{D}(T^\dagger)$ beliebig. Da nach Annahme $R_\alpha \to T^\dagger$ punktweise konvergiert, existiert für alle $\varepsilon > 0$ ein $\sigma(\varepsilon) > 0$ mit

$$\|R_{\sigma(\varepsilon)} y - T^\dagger y\|_X \leq \frac{\varepsilon}{2}.$$

Dies definiert eine monoton wachsende Funktion $\sigma : \mathbb{R}^+ \to \mathbb{R}^+$ mit $\lim_{\varepsilon \to 0} \sigma(\varepsilon) = 0$. Weiterhin ist für festes $\varepsilon > 0$ der Operator $R_{\sigma(\varepsilon)}$ stetig, und damit existiert ein $\rho(\varepsilon) > 0$ mit

$$\|R_{\sigma(\varepsilon)} z - R_{\sigma(\varepsilon)} y\|_X \leq \frac{\varepsilon}{2} \qquad \text{für alle } z \in Y \text{ mit } \|z - y\|_Y \leq \rho(\varepsilon).$$

Wieder wird dadurch eine Funktion $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ mit $\lim_{\varepsilon\to 0} \rho(\varepsilon) = 0$ definiert. Wir können dabei annehmen (durch optimale Wahl von ρ bei Wahlfreiheit), dass ρ strikt monoton und stetig ist. Nach dem Satz von der Umkehrfunktion existiert also auf $\mathcal{R}(\rho)$ eine strikt monotone und stetige Umkehrfunktion ρ^{-1} mit $\lim_{\delta\to 0} \rho^{-1}(\delta) = 0$. Wir setzen diese monoton und stetig fort auf $\mathbb{R}^+$ und definieren unsere a-priori-Strategie

$$\alpha : \mathbb{R}^+ \to \mathbb{R}^+, \qquad \delta \mapsto \sigma(\rho^{-1}(\delta)).$$

Dann gilt insbesondere $\lim_{\delta\to 0} \alpha(\delta) = 0$. Weiterhin existiert für alle $\varepsilon > 0$ ein $\delta := \rho(\varepsilon) > 0$ so, dass mit $\alpha(\delta) = \sigma(\varepsilon)$ gilt

$$\|R_{\alpha(\delta)} y^\delta - T^\dagger y\|_X \le \|R_{\sigma(\varepsilon)} y^\delta - R_{\sigma(\varepsilon)} y\|_X + \|R_{\sigma(\varepsilon)} y - T^\dagger y\|_X \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für alle $y^\delta \in B_\delta(y)$. Damit konvergiert $\|R_{\alpha(\delta)} y^\delta - T^\dagger y\|_X \to 0$ für $\delta \to 0$ und jede Familie $\{y^\delta\}_{\delta>0} \subset Y$ mit $y^\delta \in B_\delta(y)$. Also ist (R_α, α) ein konvergentes Regularisierungsverfahren. □

Wir können sogar eine Charakterisierung von a priori-Strategien angeben, die zu konvergenten Regularisierungsverfahren führen.

Satz 4.6. *Seien $T^\dagger$ nicht stetig, $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung, und $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ eine a priori-Strategie. Dann ist (R_α, α) ein Regularisierungsverfahren genau dann, wenn gilt*

(i) $\lim_{\delta\to 0} \alpha(\delta) = 0$,

(ii) $\lim_{\delta\to 0} \delta \|R_{\alpha(\delta)}\|_{\mathcal{L}(Y,X)} = 0$.

Beweis. Aus der Zerlegung (4.2) des Gesamtfehlers folgt sofort

$$\|R_{\alpha(\delta)} y^\delta - T^\dagger y\|_X \le \delta \|R_{\alpha(\delta)}\|_{\mathcal{L}(Y,X)} + \|R_{\alpha(\delta)} y - T^\dagger y\|_X \to 0 \qquad \text{für } \delta \to 0,$$

da der erste Term nach Bedingung (ii) und der zweite Term wegen der punktweisen Konvergenz von Regularisierungsoperatoren zusammen mit Bedingung (i) gegen Null geht.

Sei umgekehrt angenommen, dass entweder (i) oder (ii) nicht gilt. Ist (i) verletzt, so konvergiert $R_{\alpha(\delta)}$ nicht punktweise gegen $T^\dagger y$. Also kann (4.3) für die konstante Folge $y^\delta \equiv y$ und $\delta \to 0$ nicht gelten und damit (R_α, α) kein Regularisierungsverfahren sein. Gilt dagegen (i) aber nicht (ii), so existiert eine Nullfolge $\{\delta_n\}_{n\in\mathbb{N}}$ mit $\delta_n \|R_{\alpha(\delta_n)}\|_{\mathcal{L}(Y,X)} \ge \varepsilon > 0$ für ein $\varepsilon > 0$. Wir können also eine Folge $\{z_n\}_{n\in\mathbb{N}} \subset Y$ mit $\|z_n\|_Y = 1$ und $\delta_n \|R_{\alpha(\delta_n)} z_n\|_X \ge \varepsilon$ finden. Sei nun $y \in \mathcal{D}(T^\dagger)$ beliebig, und setze $y_n := y + \delta_n z_n$. Dann gilt $y_n \in B_{\delta_n}(y)$, aber

$$R_{\alpha(\delta_n)} y_n - T^\dagger y = (R_{\alpha(\delta_n)} y - T^\dagger y) + \delta_n R_{\alpha(\delta_n)} z_n \not\to 0,$$

da der erste Term auf der rechten Seite wegen Bedingung (i) und der punktweisen Konvergenz von R_α eine Nullfolge, aber der zweite Term nach Konstruktion keine Nullfolge ist. Also ist (4.3) verletzt und (R_α, α) daher kein Regularisierungsverfahren. Die Aussage folgt nun durch Kontraposition. □

Wegen $\|R_\alpha\|_{\mathcal{L}(Y,X)} \to \infty$ für $\alpha \to 0$ bedeutet dabei die zweite Bedingung, dass α nicht zu schnell im Verhältnis zu δ gegen Null gehen darf. Eine a priori-Strategie hat also üblicherweise die Form $\alpha(\delta) = \delta^r$ für ein $r \in (0,1)$ (wobei die Wahl von r unter anderem von der spezifischen Regularisierung $\{R_\alpha\}_{\alpha>0}$ abhängt).

A POSTERIORI-STRATEGIEN

Wie wir später sehen werden, benötigt die optimale Wahl von $\alpha(\delta)$ Informationen über die fehlerfreie (Minimum-Norm-)Lösung $x^\dagger$, die nicht leicht zugänglich sind. A posteriori-Strategien kommen dagegen ohne diese Information aus. Die Grundidee ist dabei folgende: Sei wieder $y \in \mathcal{D}(T^\dagger)$ sowie $y^\delta \in B_\delta(y)$ und betrachte das *Residuum*

$$\|TR_\alpha y^\delta - y^\delta\|_Y.$$

Gilt nun $y \in \mathcal{R}(T)$ und $\|y - y^\delta\|_Y = \delta$, so erfüllt selbst die (eigentlich gesuchte) Minimum-Norm-Lösung $x^\dagger$ wegen $Tx^\dagger = y$ nur

$$\|Tx^\dagger - y^\delta\|_Y = \|y - y^\delta\|_Y = \delta.$$

Es ist also nicht sinnvoll zu versuchen, für die Regularisierung $R_\alpha y^\delta$ ein kleineres Residuum zu erreichen. Dies motiviert das *Diskrepanzprinzip von Morozov*: Zu gegebenem $\delta > 0$ und $y^\delta \in B_\delta(y)$ wähle $\alpha = \alpha(\delta, y^\delta)$ so, dass (möglichst scharf) gilt

$$\|TR_\alpha y^\delta - y^\delta\|_Y \le \tau\delta \qquad \text{für ein } \tau > 1 \text{ unabhängig von } \delta \text{ und } y^\delta. \tag{4.4}$$

Allerdings muss dieses Prinzip nicht erfüllt sein: Ist $y \in \mathcal{R}(T)^\perp \setminus \{0\}$, so gilt selbst für exakte Daten $y^\delta = y$ und die Minimum-Norm-Lösung $x^\dagger$

$$\|Tx^\dagger - y\|_Y = \|TT^\dagger y - y\|_Y = \|P_{\overline{\mathcal{R}}} y - y\|_Y = \|y\|_Y > \tau\delta$$

für $\tau > 1$ fest und δ klein genug. Wir müssen also annehmen, dass dieser Fall nicht eintreten kann; hinreichend dafür ist, dass $\mathcal{R}(T)$ dicht in Y liegt (denn dann ist $\mathcal{R}(T)^\perp = \overline{\mathcal{R}(T)}^\perp = \{0\}$).

Für die praktische Umsetzung wählt man üblicherweise eine Nullfolge $\{\alpha_n\}_{n\in\mathbb{N}}$, berechnet sukzessive $R_{\alpha_n} y^\delta$ für $n = 1, \dots,$ und hört auf, sobald für ein α_{n^*} das Diskrepanzprinzip (4.4) erfüllt ist. Der folgende Satz liefert hierfür die Rechtfertigung.

Satz 4.7. *Sei $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung von $T^\dagger$ mit $\mathcal{R}(T)$ dicht in Y, $\{\alpha_n\}_{n\in\mathbb{N}}$ eine streng monoton fallende Nullfolge, und $\tau > 1$. Ist die Familie $\{TR_\alpha\}_{\alpha>0}$ gleichmäßig beschränkt, so existiert für alle $y \in \mathcal{D}(T^\dagger)$, $\delta > 0$, und $y^\delta \in B_\delta(y)$ ein $n^* \in \mathbb{N}$, so dass gilt*

$$\|TR_{\alpha_{n^*}} y^\delta - y^\delta\|_Y \le \tau\delta < \|TR_{\alpha_n} y^\delta - y^\delta\|_Y \qquad \textit{für alle } n < n^*. \tag{4.5}$$

Beweis. Wir gehen vor wie im Beweis von Satz 4.3. Auf $\mathcal{D}(T^\dagger)$ konvergiert TR_α punktweise gegen $TT^\dagger = P_{\overline{\mathcal{R}}}$ und damit wegen der gleichmäßigen Beschränktheit auf ganz $Y = \overline{\mathcal{D}(T^\dagger)}$. Für alle $y \in \mathcal{D}(T^\dagger) = \mathcal{R}(T)$ und $y^\delta \in B_\delta(y)$ folgt daraus

$$\lim_{n\to\infty} \|TR_{\alpha_n} y^\delta - y^\delta\|_Y = \|P_{\overline{\mathcal{R}}} y^\delta - y^\delta\|_Y = 0$$

wegen $\overline{\mathcal{R}(T)} = Y$ und damit die Behauptung. □

Um zu zeigen, dass das Diskrepanzprinzip tatsächlich zu einem Regularisierungsverfahren führt, muss man es in Kombination mit einer konkreten Regularisierung betrachten. Wir werden dies in den nächsten Kapiteln nachholen.

HEURISTISCHE STRATEGIEN

Heuristische Strategien kommen sogar ohne Kenntnis des Fehlerniveaus δ aus. Dies ist in der Praxis sehr relevant, denn oft hat man keine hinreichend scharfe Abschätzung für δ. Das folgende einschneidende Resultat – in der Literatur unter dem Namen *Bakushinskiĭ-Veto* bekannt, siehe [Bakushinskiĭ 1985] – sagt jedoch, dass das im Allgemeinen nicht klappen kann.

Satz 4.8. *Sei $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung von $T^\dagger$. Existiert eine heuristische Parameterwahlstrategie α so, dass (R_α, α) ein Regularisierungsverfahren ist, dann ist $T^\dagger$ stetig.*

Beweis. Angenommen, es gäbe solch eine Parameterwahlstrategie $\alpha : Y \to \mathbb{R}^+$. Dann können wir die Abbildung

$$R : Y \to X, \qquad y \mapsto R_{\alpha(y)} y,$$

definieren (die erstmal nicht linear sein muss). Sei nun $y \in \mathcal{D}(T^\dagger)$ beliebig und betrachte eine beliebige Folge $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T^\dagger)$ mit $y_n \to y$. Dann ist einerseits natürlich $y_n \in B_\delta(y_n)$ für alle $\delta > 0$ und $n \in \mathbb{N}$, und aus der Annahme (4.3) für $y^\delta = y = y_n$ fest und $\delta \to 0$ folgt, dass $Ry_n = T^\dagger y_n$ für alle $n \in \mathbb{N}$ (und damit R tatsächlich auf $\mathcal{D}(T^\dagger)$ linear) ist. Außerdem ist für $\delta_n := \|y_n - y\|_Y$ auch $y_n \in B_{\delta_n}(y)$, und aus (4.3) folgt mit $n \to \infty$

$$T^\dagger y_n = Ry_n = R_{\alpha(y_n)} y_n \to T^\dagger y,$$

d. h. $T^\dagger$ ist stetig auf $\mathcal{D}(T^\dagger)$. □

Insbesondere kann für kompakte Operatoren mit unendlichdimensionalem Bild keine heuristische Parameterwahlstrategie zu einem konvergenten Regularisierungsverfahren führen. Natürlich heißt das nicht, dass solche Strategien in der Praxis nicht eingesetzt werden. Zum einen verbietet das Veto keine Strategien für endlichdimensionale schlecht

gestellte Probleme (wie etwa sehr schlecht konditionierte lineare Gleichungssysteme); diese sind dann allerdings notwendigerweise dimensionsabhängig. Schaut man sich den Beweis außerdem genau an, erkennt man auch, dass der zentrale Schritt darin besteht, die Parameterwahlstrategie auf Daten $y^\delta \in \mathcal{D}(T^\dagger)$ anzuwenden. Der schlimmstmögliche Fall für die gestörten Daten ist also $y^\delta \in \mathcal{R}(T)$ (wegen $\mathcal{R}(T)^\perp = \mathcal{N}(T^\dagger)$ spielt nur dieser Teilraum von $\mathcal{D}(T^\dagger)$ eine Rolle), und in diesem Fall kann keine Konvergenz garantiert werden. Nun ist üblicherweise in der Praxis T ein kompakter (d. h. glättender) Operator, während Störungen eher zufälligen Charakter haben und damit nicht in $\mathcal{R}(T)$ liegen. Im „üblichen“ Fall kann ein heuristisches Verfahren also sehr wohl funktionieren. In der Tat kann man zeigen, dass unter der zusätzlichen Annahme $y^\delta \notin \mathcal{D}(T^\dagger)$ eine ganze Klasse von beliebten heuristischen Strategien zu einem Regularisierungsverfahren führen. Auch hierfür muss man die Kombination mit konkreten Regularierungen betrachten, wir geben aber bereits ein paar Beispiele an:

(i) Das *Quasioptimalitäts-Prinzip* wählt eine endliche streng monoton fallende Folge $\{\alpha_n\}_{n\in\{1,\dots,N\}}$ und bestimmt $\alpha(y^\delta)$ als α_{n^*} mit

$$n^* \in \arg\min_{1\le n<N} \|R_{\alpha_{n+1}} y^\delta - R_{\alpha_n} y^\delta\|_X. \tag{4.6}$$

(ii) Die *Hanke–Raus-Regel* bestimmt

$$\alpha(y^\delta) \in \arg\min_{\alpha>0} \frac{1}{\sqrt{\alpha}} \|TR_\alpha y^\delta - y^\delta\|_Y. \tag{4.7}$$

(iii) Das *L-Kurven-Kriterium*[1] bestimmt

$$\alpha(y^\delta) \in \arg\min_{\alpha>0} \|R_\alpha y^\delta\|_X \|TR_\alpha y^\delta - y^\delta\|_Y. \tag{4.8}$$

Alle diese Verfahren basieren in der einen oder anderen Weise darauf, aus dem Residuum eine möglichst gute Schätzung des Fehlerniveaus zu gewinnen, die dann analog zu a priori- oder a posteriori-Strategien eingesetzt werden kann. Einen umfassenden Vergleich dieser und weiterer Strategien findet man in [Bauer & Lukas 2011].

4.2 KONVERGENZRATEN

Eine wesentliche Fragestellung in der Regularisierung inverser Probleme ist das Herleiten von Fehlerabschätzungen der Form

$$\|R_{\alpha(\delta,y^\delta)} y^\delta - T^\dagger y\|_X \le \varphi(\delta)$$

[1] Der Name stammt aus der Praxis: Zeichnet man sich die Kurve $\alpha \mapsto (\|TR_\alpha y^\delta - y^\delta\|_Y, \|R_\alpha y^\delta\|_X)$ (bzw. ausgewählte Punkte auf dieser) in einer doppelt-logarithmischen Darstellung auf, so hat diese – oft mehr oder weniger – die Form eines “L”; der gesuchte Parameter ist dann derjenige, der in dessen “Knick” liegt.

für eine monotone Funktion $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ mit $\lim_{t\to 0} \varphi(t) = 0$. Wir sind insbesondere an Abschätzungen für den *schlimmstmöglichen Fehler*

$$\mathcal{E}(y,\delta) := \sup_{y^\delta \in B_\delta(y)} \|R_{\alpha(\delta,y^\delta)} y^\delta - T^\dagger y\|_X \tag{4.9}$$

(der für konvergente Regularisierungsverfahren nach (4.3) gegen Null geht für $\delta \to 0$ und alle $y \in \mathcal{D}(T^\dagger)$) interessiert. Dabei muss φ in irgendeiner Form von y abhängen, denn sonst könnte man a priori – ohne Kenntnis von y und y^δ – garantierte Fehlerschranken für Regularisierungsverfahren angeben. Da die Konvergenz von $R_\alpha \to T^\dagger$ lediglich punktweise aber nicht gleichmäßig ist, kann man solche Abschätzungen jedoch nicht erwarten.

Satz 4.9. *Sei (R_α, α) ein Regularisierungsverfahren. Existiert eine Funktion $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ mit $\lim_{t\to 0} \varphi(t) = 0$ und*

$$\sup_{y \in \mathcal{D}(T^\dagger) \cap B_Y} \mathcal{E}(y,\delta) \le \varphi(\delta), \tag{4.10}$$

so ist $T^\dagger$ stetig.

Beweis. Sei $y \in \mathcal{D}(T^\dagger) \cap B_Y$ und $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T^\dagger) \cap B_Y$ eine Folge mit $y_n \to y$. Mit $\delta_n := \|y - y_n\|_Y \to 0$ für $n \to \infty$ gilt dann

$$\begin{aligned}
\|T^\dagger y_n - T^\dagger y\|_X &\le \|T^\dagger y_n - R_{\alpha(\delta_n,y_n)} y_n\|_X + \|R_{\alpha(\delta_n,y_n)} y_n - T^\dagger y\|_X \\
&\le \mathcal{E}(y_n,\delta_n) + \mathcal{E}(y,\delta_n) \\
&\le 2\varphi(\delta_n) \to 0.
\end{aligned}$$

Also ist $T^\dagger$ stetig auf $\mathcal{D}(T^\dagger) \cap B_Y$, und wegen der Linearität von $T^\dagger$ muss daher $T^\dagger$ stetig auf ganz $\mathcal{D}(T^\dagger)$ sein. □

Die Konvergenz kann also beliebig langsam sein; Kenntnis von δ alleine reicht daher nicht aus, um Fehlerschranken angeben zu können – wir brauchen weitere Annahmen an die exakten Daten y bzw. die gesuchte Minimum-Norm-Lösung $x^\dagger = T^\dagger y$. Anhand von Satz 4.9 sieht man, dass die Existenz von Konvergenzraten eng mit der Stetigkeit von $T^\dagger$ auf abgeschlossenen Teilmengen verknüpft ist. Wir betrachten daher für $\mathcal{M} \subset X$ und $\delta > 0$ die Größe

$$\varepsilon(\mathcal{M},\delta) := \sup\{\|x\|_X : x \in \mathcal{M},\ \|Tx\|_Y \le \delta\},$$

die man als *bedingtes Stetigkeitsmodul* von $T^\dagger : \mathcal{R}(T) \cap \delta B_Y \to \mathcal{M}$ interpretieren kann. Das Stetigkeitsmodul gibt nun eine untere Schranke für den schlimmstmöglichen Fehler an. Da sowohl $\varepsilon(\mathcal{M},\delta)$ als auch $\mathcal{E}(y,\delta)$ unendlich sind falls $\mathcal{M} \cap \mathcal{N}(T) \neq \{0\}$ und $\mathcal{M}$ unbeschränkt ist, betrachten wir in Folge den interessanteren Fall $\mathcal{M} \subset \mathcal{N}(T)^\perp$.

Satz 4.10. *Sei (R_α, α) ein Regularisierungsverfahren. Dann gilt für alle $\delta > 0$ und $\mathcal{M} \subset \mathcal{N}(T)^\perp$*

$$\sup_{y \in \mathcal{D}(T^\dagger), T^\dagger y \in \mathcal{M}} \mathcal{E}(y, \delta) \geq \varepsilon(\mathcal{M}, \delta).$$

Beweis. Sei $x \in \mathcal{M}$ mit $\|Tx\|_Y \leq \delta$. Für $y^\delta = 0$ erhalten wir dann wegen $x \in \mathcal{N}(T)^\perp$

$$\|x\|_X = \|T^\dagger Tx - R_{\alpha(\delta,0)} 0\|_X \leq \mathcal{E}(Tx, \delta)$$

und damit

$$\varepsilon(\mathcal{M}, \delta) = \sup_{x \in \mathcal{M}, \|Tx\|_Y \leq \delta} \|x\|_X \leq \sup_{x \in \mathcal{M}, \|Tx\|_Y \leq \delta} \mathcal{E}(Tx, \delta) \leq \sup_{T^\dagger y \in \mathcal{M}, y \in \mathcal{D}(T^\dagger)} \mathcal{E}(y, \delta),$$

da $\mathcal{D}(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ und $\mathcal{R}(T)^\perp = \mathcal{N}(T^\dagger)$. □

Für eine geeignete Wahl von $\mathcal{M}$ kann man nun scharfe Schranken für $\varepsilon(\mathcal{M}, \delta)$ angeben. Wir betrachten hier für kompakte Operatoren $K \in \mathcal{K}(X, Y)$ Mengen der Form

$$X_{\nu,\rho} = \{|K|^\nu w \in X : \|w\|_X \leq \rho\} \subset \mathcal{R}(|K|^\nu).$$

Wegen der Definition von $|K|^\nu w$ über die Spektralzerlegung von K gilt insbesondere $X_{\nu,\rho} \subset \overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$.

Satz 4.11. *Sei $K \in \mathcal{K}(X, Y)$ und $\nu, \rho > 0$. Dann gilt für alle $\delta > 0$*

$$\varepsilon(X_{\nu,\rho}, \delta) \leq \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}.$$

Beweis. Sei $x \in X_{\nu,\rho}$ und $\|Kx\|_Y \leq \delta$. Dann existiert $w \in X$ mit $x = |K|^\nu w$ und $\|w\|_X \leq \rho$. Aus der Interpolationsungleichung in Lemma 3.14 mit $s = \nu$ und $r = \nu + 1$ sowie den Eigenschaften aus Lemma 3.13 folgt daher

$$\begin{aligned} \|x\|_X = \||K|^\nu w\|_X &\leq \||K|^{\nu+1} w\|_X^{\frac{\nu}{\nu+1}} \|w\|_X^{\frac{1}{\nu+1}} = \|K|K|^\nu w\|_Y^{\frac{\nu}{\nu+1}} \|w\|_X^{\frac{1}{\nu+1}} \\ &= \|Kx\|_Y^{\frac{\nu}{\nu+1}} \|w\|_X^{\frac{1}{\nu+1}} \leq \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}. \end{aligned}$$

Supremum über alle $x \in X_{\nu,\rho}$ mit $\|Kx\|_Y \leq \delta$ ergibt die Aussage. □

Dies ist zwar nur eine obere Schranke, es gibt jedoch immer zumindest eine Folge, für die sie angenommen wird.

Satz 4.12. *Sei $K \in \mathcal{K}(X, Y)$ und $\nu, \rho > 0$. Dann existiert eine Nullfolge $\{\delta_n\}_{n \in \mathbb{N}}$ mit*

$$\varepsilon(X_{\nu,\rho}, \delta_n) = \delta_n^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}.$$

Beweis. Sei $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System für K, und setze $\delta_n := \rho\sigma_n^{\nu+1}$ und $x_n := |K|^\nu(\rho v_n)$. Da die Singulärwerte eine Nullfolge bilden, gilt $\delta_n \to 0$. Weiter ist nach Konstruktion $x_n \in X_{\nu,\rho}$. Aus $\sigma_n = (\rho^{-1}\delta_n)^{\frac{1}{\nu+1}}$ folgt nun

$$x_n = \rho|K|^\nu v_n = \rho\sigma_n^\nu v_n = \delta_n^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}v_n,$$

da σ_n^ν Eigenwert von $|K|^\nu$ mit Eigenvektor v_n ist. Also ist $\|x_n\|_X = \delta_n^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}$. Analog folgt

$$K^*Kx_n = \delta_n^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}\sigma_n^2 v_n = \delta_n^{\frac{\nu+2}{\nu+1}}\rho^{-\frac{1}{\nu+1}}v_n$$

und damit

$$\|Kx_n\|_Y^2 = (Kx_n, Kx_n)_Y = (K^*Kx_n, x_n)_X = \delta_n^2.$$

Es gilt daher für alle $n \in \mathbb{N}$

$$\varepsilon(X_{\nu,\rho}, \delta_n) = \sup_{x\in X_{\nu,\rho},\ \|Kx\|_Y\le\delta_n} \|x\|_X \ge \|x_n\|_X = \delta_n^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}},$$

woraus zusammen mit Satz 4.11 die Gleichheit folgt. □

Hieraus folgt, dass für einen kompakten Operator K mit unendlichdimensionalem Bild kein Regularisierungsverfahren einen Gesamtfehler ergeben kann, der für $\delta \to 0$ garantiert schneller als $\delta_n^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}$ gegen Null geht – und das auch nur unter der Zusatzannahme $x^\dagger \in X_{\nu,\rho}$.[2] Insbesondere konvergiert der Gesamtfehler also stets langsamer als der Datenfehler.

Da das bedingte Stetigkeitsmodul nur eine untere Schranke ist, geben wir noch etwas Luft und nennen ein Regularisierungsverfahren daher *ordnungsoptimal* (für ν und ρ), falls eine Konstante $c = c(\nu) \ge 1$ existiert mit

$$\mathcal{E}(Kx^\dagger, \delta) \le c\delta^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}} \qquad \text{für alle } x^\dagger \in X_{\nu,\rho}. \tag{4.11}$$

Lässt man zu, dass die Konstante von $x^\dagger$ abhängt – ist man also nur an *Konvergenzraten* interessiert – so betrachtet man

$$X_\nu := \bigcup_{\rho>0} X_{\nu,\rho} = \mathcal{R}(|K|^\nu)$$

und bezeichnet ein Regularisierungsverfahren als ordnungsoptimal für ν, falls ein $c = c(x^\dagger) \ge 1$ existiert mit

$$\mathcal{E}(Kx^\dagger, \delta) \le c\delta^{\frac{\nu}{\nu+1}} \qquad \text{für alle } x^\dagger \in X_\nu.$$

Die Bedingung $x^\dagger \in X_{\nu,\rho}$ bezeichnet man dabei als *Quellbedingung*, das Element $w \in X$ mit $|K|^\nu w = x^\dagger$ als *Quelldarstellung*. Da K ein kompakter (d. h. glättender) Operator ist, stellen

[2] Dies ist ein zentrales Paradigma in der Theorie – und Praxis – der inversen Probleme: *Stabilität ist nur unter Zusatzannahmen erreichbar!*

Quellbedingungen abstrakte Glattheitsbedingungen dar. Für den Integrationsoperator K aus Beispiel 3.11 bedeutet etwa $x \in X_{2,\rho}$, dass $x = K^*Kw = \int_t^1 \int_0^s w(r)\,dr\,ds$ eine zweite Ableitung w (im L^2-Sinn) mit Norm kleiner oder gleich ρ besitzt. Mit Hilfe der Singulärwertzerlegung sieht man leicht, dass die Bedingung $x^\dagger \in X_\nu$ einer verschärften Picard-Bedingung entspricht, d. h. einer Abklingrate der Fourier-Koeffizienten von y, die umso schneller im Vergleich zu den Singulärwerten ist, desto grösser ν ist.

Lemma 4.13. *Sei $K \in \mathcal{K}(X,Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ und sei $y \in \mathcal{R}(K)$. Dann ist $x^\dagger = K^\dagger y \in X_\nu$ genau dann, wenn gilt*

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2-2\nu} |\,(y,u_n)_Y\,|^2 < \infty. \tag{4.12}$$

Beweis. Nach Definition und der Darstellung (3.14) ist $K^\dagger y \in X_\nu$ genau dann, wenn ein $w \in X$ existiert mit

$$\sum_{n\in\mathbb{N}} \sigma_n^{-1} (y,u_n)_Y\, v_n = K^\dagger y = |K|^\nu w = \sum_{n\in\mathbb{N}} \sigma_n^\nu (w,v_n)_X\, v_n.$$

Da die v_n ein Orthonormalsystem bilden, können wir einen Koeffizientenvergleich durchführen und erhalten

$$\sigma_n^{-1} (y,u_n)_Y = \sigma_n^\nu (w,v_n)_X \qquad \text{für alle } n \in \mathbb{N}.$$

Wie im Beweis von Satz 3.10 folgt nun, dass $w \in X$ dann und nur dann gilt, wenn die Reihe $\sum_{n\in\mathbb{N}} |\,(w,v_n)_X\,|^2$ konvergiert, d. h. (4.12) gilt. □

Tatsächlich impliziert die Ordnungsoptimalität bereits die Regularisierungseigenschaft eines Verfahrens. Dies ist nützlich, denn es ist manchmal leichter, die Optimalität nachzuweisen als die Regularisierungseigenschaft. Dabei müssen wir berücksichtigen, dass die optimale Parameterwahl von der Quellbedingung abhängen kann.

Satz 4.14. *Seien $K \in \mathcal{K}(X,Y)$ mit $\mathcal{R}(K)$ dicht in Y, $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung und $\alpha(\delta, y^\delta)$ eine Parameterwahlstrategie. Existiert ein $\tau_0 \geq 1$ so dass R_α zusammen mit $\alpha_\tau := \alpha(\tau\delta, y^\delta)$ für alle $\tau > \tau_0$ die Bedingung (4.11) für ein $\nu > 0$ und alle $\rho > 0$ erfüllt, so ist (R_α, α_τ) für alle $\tau > \tau_0$ ein Regularisierungsverfahren.*

Beweis. Wir müssen also zeigen, dass aus der *gleichmäßigen* Konvergenz des schlimmstmöglichen Fehlers für alle $x^\dagger \in X_{\nu,\rho}$ die *punktweise* Konvergenz für alle $x^\dagger \in \mathcal{R}(K^\dagger)$ folgt. Wir konstruieren dafür ein geeignetes $x_\varepsilon \in X_{\nu,\rho}$, das wir in die Fehlerabschätzung einschieben, und wenden dann die Ordnungsoptimalität an. Die Konstante τ wird dabei verwendet, um das Fehlerniveau – und damit die Parameterwahl – bezüglich $Kx^\dagger$ auf Kx_ε anwenden zu können.

Sei also $y \in \mathcal{D}(K^\dagger) = \mathcal{R}(K)$ und sei $x^\dagger = K^\dagger y$ (und damit $Kx^\dagger = y$). Sei weiterhin $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K. Für beliebiges $\varepsilon > 0$ wählen wir nun ein $N_\varepsilon \in \mathbb{N}$ so, dass $\sigma_{N_\varepsilon} \geq \varepsilon > \sigma_{N_\varepsilon+1}$ gilt und setzen

$$x_\varepsilon := \sum_{n=1}^{N_\varepsilon} \left(x^\dagger, v_n\right)_X v_n$$

sowie

$$\begin{aligned} y_\varepsilon := Kx_\varepsilon &= \sum_{n=1}^{N_\varepsilon} \left(x^\dagger, v_n\right)_X Kv_n = \sum_{n=1}^{N_\varepsilon} \left(x^\dagger, v_n\right)_X \sigma_n u_n \\ &= \sum_{n=1}^{N_\varepsilon} \left(x^\dagger, K^* u_n\right)_X u_n = \sum_{n=1}^{N_\varepsilon} (y, u_n)_X \, u_n. \end{aligned}$$

Da $\{u_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K)}$ und $\{v_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$ ist, besitzen $x^\dagger = K^\dagger y \in \mathcal{N}(K)^\perp$ und $y = Kx^\dagger \in \mathcal{R}(K)$ die Reihendarstellungen

$$x^\dagger = \sum_{n\in\mathbb{N}} \left(x^\dagger, v_n\right)_X v_n, \qquad y = \sum_{n\in\mathbb{N}} (y, u_n)_Y \, u_n.$$

Damit ist

$$\|x^\dagger - x_\varepsilon\|_X^2 = \sum_{n=N_\varepsilon+1}^{\infty} \left|\left(x^\dagger, v_n\right)_X\right|^2$$

und

$$\begin{aligned} \|y - y_\varepsilon\|_Y^2 = \sum_{n=N_\varepsilon+1}^{\infty} \left|(y, u_n)_Y\right|^2 &= \sum_{n=N_\varepsilon+1}^{\infty} \sigma_n^2 \left|\left(x^\dagger, v_n\right)_X\right|^2 \\ &< \varepsilon^2 \sum_{n=N_\varepsilon+1}^{\infty} \left|\left(x^\dagger, v_n\right)_X\right|^2 = \varepsilon^2 \|x^\dagger - x_\varepsilon\|_X^2, \end{aligned} \tag{4.13}$$

nach Wahl von N_ε. Insbesondere konvergieren $x_\varepsilon \longrightarrow x^\dagger$ und $y_\varepsilon \longrightarrow y$ mit $\varepsilon \longrightarrow 0$ (und zwar monoton).

Nach Konstruktion ist $y_\varepsilon \in \mathcal{R}(K)$ und $x_\varepsilon \in \mathcal{N}(K)^\perp$, und damit gilt $x_\varepsilon = K^\dagger y_\varepsilon$. Nach Lemma 4.13 ist deshalb $x_\varepsilon \in X_\nu$ für alle $\nu > 0$, denn wegen $(y_\varepsilon, u_n)_Y = 0$ für $n > N_\varepsilon$ ist die Reihe in (4.12) endlich. Es existiert also ein $w_\varepsilon \in X$ mit $x_\varepsilon = |K|^\nu w_\varepsilon$, d. h.

$$\sum_{n=1}^{N_\varepsilon} \left(x^\dagger, v_n\right)_X v_n = x_\varepsilon = |K|^\nu w_\varepsilon = \sum_{n\in\mathbb{N}} \sigma_n^\nu \, (w_\varepsilon, v_n)_X \, v_n.$$

Da $\mathcal{R}(K)$ dicht in Y ist, kann K kein endlichdimensionales Bild haben, was $\sigma_n > 0$ für alle $n \in \mathbb{N}$ impliziert. Aus der Orthonormalität der v_n folgt daher

$$(w_\varepsilon, v_n)_X = \begin{cases} \sigma_n^{-\nu} \left(x^\dagger, v_n\right)_X & n \leq N_\varepsilon, \\ 0 & n > N_\varepsilon. \end{cases}$$

Deshalb gilt

$$\|w_\varepsilon\|_X^2 = \sum_{n=1}^{N_\varepsilon} |(w_\varepsilon, v_n)_X|^2 = \sum_{n=1}^{N_\varepsilon} \sigma_n^{-2\nu} \left| \left(x^\dagger, v_n\right)_X \right|^2 \leq \varepsilon^{-2\nu} \sum_{n\in\mathbb{N}} \left| \left(x^\dagger, v_n\right)_X \right|^2 = \varepsilon^{-2\nu} \|x^\dagger\|_X^2, \tag{4.14}$$

wieder nach Wahl von N_ε, und damit $x_\varepsilon \in X_{\nu,\rho}$ mit $\rho = \varepsilon^{-\nu}\|x^\dagger\|_X$.

Seien nun $y^\delta \in B_\delta(y)$ und $\tau > \tau_0 \geq 1$. Wir wählen nun

$$\varepsilon(\delta) := \inf\left\{\varepsilon > 0 : \|y - y_\varepsilon\|_Y \geq \frac{\tau - \tau_0}{\tau + \tau_0}\delta\right\}.$$

Nach Definition und wegen der Linksstetigkeit (aufgrund der Wahl von N_ε) und Monotonie von $\varepsilon \mapsto \|y - y_\varepsilon\|_Y$ gilt dann insbesondere

$$\|y - y_{\varepsilon(\delta)}\|_Y \leq \frac{\tau - \tau_0}{\tau + \tau_0}\delta \leq \|y - y_{2\varepsilon(\delta)}\|_Y \tag{4.15}$$

und damit

$$\|y^\delta - y_{\varepsilon(\delta)}\|_Y \leq \|y^\delta - y\|_Y + \|y - y_{\varepsilon(\delta)}\|_Y \leq \left(1 + \frac{\tau - \tau_0}{\tau + \tau_0}\right)\delta =: \tilde{\delta}.$$

Ist y^δ also ein gestörter Messwert zu den exakten Daten y mit Fehlerniveau δ, so ist y^δ auch ein gestörter Messwert zu $y_{\varepsilon(\delta)}$ mit Fehlerniveau $\tilde{\delta}$. Für $\tilde{\tau} := \frac{1}{2}(\tau + \tau_0) > \tau_0$ ist dann $\tilde{\tau}\tilde{\delta} = \tau\delta$ und damit

$$\alpha_{\tilde{\tau}}(\tilde{\delta}, y^\delta) = \alpha(\tilde{\tau}\tilde{\delta}, y^\delta) = \alpha(\tau\delta, y^\delta) = \alpha_\tau(\delta, y^\delta),$$

d. h. die Parameterwahlstrategien α_τ bezüglich y und $\alpha_{\tilde{\tau}}$ bezüglich $y_{\varepsilon(\delta)}$ stimmen für festes y^δ überein. Aus der Ordnungsoptimalität (4.11) von $(R_\alpha, \alpha_{\tilde{\tau}})$ folgt daher für $x_\varepsilon \in X_{\nu,\rho}$, dass gilt

$$\begin{aligned}\|R_{\alpha_\tau(\delta,y^\delta)}y^\delta - x_\varepsilon\|_X = \|R_{\alpha_{\tilde{\tau}}(\tilde{\delta},y^\delta)}y^\delta - K^\dagger y_\varepsilon\|_X \leq \mathcal{E}(y_\varepsilon, \tilde{\delta}) &\leq c\tilde{\delta}^{\frac{\nu}{\nu+1}} \left(\varepsilon^{-\nu}\|x^\dagger\|_X\right)^{\frac{1}{\nu+1}} \\ &=: c_{\tau,\nu}\left(\frac{\delta}{\varepsilon}\right)^{\frac{\nu}{\nu+1}} \|x^\dagger\|_X^{\frac{1}{\nu+1}}.\end{aligned}$$

Damit gilt

$$\begin{aligned}\|R_{\alpha_\tau(\delta,y^\delta)}y^\delta - x^\dagger\|_X &\leq \|R_{\alpha_\tau(\delta,y^\delta)}y^\delta - x_{\varepsilon(\delta)}\|_X + \|x_{\varepsilon(\delta)} - x^\dagger\|_X \\ &\leq c_{\tau,\nu}\left(\frac{\delta}{\varepsilon(\delta)}\right)^{\frac{\nu}{\nu+1}} \left(\|x^\dagger\|_X\right)^{\frac{1}{\nu+1}} + \|x_{\varepsilon(\delta)} - x^\dagger\|_X,\end{aligned}$$

und es bleibt zu zeigen, dass sowohl $\delta\varepsilon(\delta)^{-1} \to 0$ als auch $x_{\varepsilon(\delta)} \to x^\dagger$ für $\delta \to 0$ gehen. Da $\varepsilon(\delta) > 0$ monoton fallend für $\delta \to 0$ und deshalb konvergent ist, müssen wir nur zwei Fälle unterscheiden:

(i) $\varepsilon(\delta) \to \varepsilon_0 > 0$ für $\delta \to 0$. In diesem Fall ist offensichtlich $\delta\varepsilon(\delta)^{-1} \le \delta\varepsilon_0{}^{-1} \to 0$. Aus (4.15) folgt weiter

$$\|y - y_{\varepsilon_0}\|_Y = \lim_{\delta\to 0} \|y - y_{\varepsilon(\delta)}\|_X \le \lim_{\delta\to 0} \frac{\tau - \tau_0}{\tau + \tau_0}\delta = 0$$

und $x_{\varepsilon_0} = K^\dagger y_{\varepsilon_0} = K^\dagger y = x^\dagger$.

(ii) $\varepsilon(\delta) \to 0$ für $\delta \to 0$. Dann folgt sofort $x_{\varepsilon(\delta)} \to x^\dagger$. Wegen (4.15) und (4.13) gilt weiter

$$\frac{\tau - \tau_0}{\tau + \tau_0}\delta \le \|y - y_{2\varepsilon(\delta)}\|_Y \le 2\varepsilon(\delta)\|x^\dagger - x_{2\varepsilon(\delta)}\|_X$$

und damit

$$\frac{\delta}{\varepsilon(\delta)} \le 2\frac{\tau + \tau_0}{\tau - \tau_0}\|x^\dagger - x_{2\varepsilon(\delta)}\|_X \to 0.$$

Also gilt $R_{\alpha_\tau(\delta,y^\delta)}y^\delta \to x^\dagger$ für alle $y \in \mathcal{D}(K^\dagger)$ und $y^\delta \in B_\delta(y)$, und (R_α, α_τ) ist ein Regularisierungsverfahren. □

Schließlich sei noch erwähnt, dass es auch schwächere Quellbedingungen mit allgemeineren *Indexfunktionen* φ als Potenzen von K^*K gibt. Ein Beispiel sind *logarithmische Quellbedingungen* der Form $x^\dagger \in \mathcal{R}(-\ln|K|)$, die für exponentiell schlecht gestellte Probleme günstiger sind; siehe z. B. [Hohage 2000]. Tatsächlich kann man zeigen, dass für jedes $x^\dagger \in X$ eine Indexfunktion φ mit $x^\dagger \in \mathcal{R}(\varphi(|K|))$ existiert; siehe [Mathé & Hofmann 2008].

5 SPEKTRALE REGULARISIERUNG

Wir haben gesehen, dass die Regularisierung einer schlechtgestellten Operatorgleichung $Tx = y$ darin besteht, die (unbeschränkte) Moore–Penrose-Inverse $T^\dagger$ durch eine Familie $\{R_\alpha\}_{\alpha>0}$ von Operatoren zu ersetzen, die für jedes $\alpha > 0$ stetig auf ganz Y sind und für $\alpha \to 0$ punktweise auf $\mathcal{D}(T^\dagger)$ gegen $T^\dagger$ konvergieren. Für kompakte Operatoren $K \in \mathcal{K}(X, Y)$ lassen sich solche Regularisierungen mit Hilfe der Singulärwertzerlegung konstruieren. Dafür verwenden wir, dass nach Folgerung 3.6 für $y \in \mathcal{D}(K^\dagger)$ gilt

$$K^\dagger y = (K^*K)^\dagger K^* y.$$

Sei nun $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K. Dann ist nach Konstruktion insbesondere $\{(\sigma_n^2, v_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K^*K, und wir können nach Satz 3.10 schreiben

$$\begin{aligned}(K^*K)^\dagger K^* y &= \sum_{n\in\mathbb{N}} \sigma_n^{-2} \left(K^* y, v_n\right)_X v_n = \sum_{n\in\mathbb{N}} \sigma_n^{-2}\sigma_n \left(y, u_n\right)_Y v_n \\ &= \sum_{n\in\mathbb{N}} \varphi(\sigma_n^2)\sigma_n \left(y, u_n\right)_Y v_n\end{aligned}$$

mit $\varphi(\lambda) = \lambda^{-1}$. Die Unstetigkeit von $K^\dagger$ rührt nun daher, dass φ auf $(0, \|K^*K\|_{\mathcal{L}(X,X)}]$ unbeschränkt und $\{\sigma_n\}_{n\in\mathbb{N}}$ eine Nullfolge ist. Um zu regularisieren, ersetzen wir deshalb φ durch eine Familie $\{\varphi_\alpha\}_{\alpha>0}$ von *beschränkten* Funktionen, die punktweise gegen φ konvergieren. Wir schreiben hier und in Folge kurz $\kappa := \|K\|^2_{\mathcal{L}(X,Y)} = \|K^*K\|_{\mathcal{L}(X,X)}$.

Definition 5.1. Sei $\{\varphi_\alpha\}_{\alpha>0}$ eine Familie von stückweise stetigen und beschränkten Funktionen $\varphi_\alpha : [0, \kappa] \to \mathbb{R}$. Gilt

(i) $\lim\limits_{\alpha\to 0} \varphi_\alpha(\lambda) = \dfrac{1}{\lambda}$ für alle $\lambda \in (0, \kappa]$ und

(ii) $\lambda|\varphi_\alpha(\lambda)| \le C_\varphi$ für ein $C_\varphi > 0$ und alle $\lambda \in (0, \kappa]$ und $\alpha > 0$,

so heißt $\{\varphi_\alpha\}_{\alpha>0}$ *(regularisierender) Filter.*

Beachten Sie, dass die Filtereigenschaft einzig über die Norm von K abhängt – insbesondere gilt: sind die Eigenschaften für alle $\lambda > 0$ erfüllt, ist $\{\varphi_\alpha\}_{\alpha>0}$ ein Filter für *beliebige* kompakte Operatoren.

Die Idee ist nun, als Regularisierungsoperator $R_\alpha := \varphi_\alpha(K^*K)K^*$ zu wählen, d. h. für $y \in Y$

$$\begin{aligned} R_\alpha y = \varphi_\alpha(K^*K)K^*y &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\,(K^*y, v_n)_Y\, v_n + \varphi_\alpha(0)P_{\mathcal{N}}K^*y \\ &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n\,(y, u_n)_Y\, v_n, \end{aligned}$$

wegen $K^*y \in \overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$. (Im Gegensatz zum Filter hängt die Regularisierung natürlich sehr wohl über das singuläre System von K ab.)

Beispiel 5.2. (i) Die *abgeschnittene Singulärwertzerlegung* entsteht durch die Wahl

$$\varphi_\alpha(\lambda) = \begin{cases} \frac{1}{\lambda} & \text{falls } \lambda \geq \alpha, \\ 0 & \text{sonst.} \end{cases} \tag{5.1}$$

Offensichtlich ist φ_α beschränkt (durch $\frac{1}{\alpha}$) und stückweise stetig, konvergiert für $\lambda > 0$ und $\alpha \to 0$ gegen $\frac{1}{\lambda}$, und erfüllt die Beschränktheitsbedingung für $C_\varphi = 1$. Der zugehörige Regularisierungsoperator ist gegeben durch

$$R_\alpha y = \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n\,(y, u_n)_Y\, v_n = \sum_{\sigma_n \geq \sqrt{\alpha}} \frac{1}{\sigma_n}\,(y, u_n)_Y\, v_n, \tag{5.2}$$

woraus sich auch der Name ergibt.

(ii) Die *Tikhonov-Regularisierung* entsteht durch die Wahl

$$\varphi_\alpha(\lambda) = \frac{1}{\lambda + \alpha}.$$

Wiederum ist φ_α beschränkt (durch $\frac{1}{\alpha}$) und stetig, konvergiert für $\lambda > 0$ und $\alpha \to 0$ gegen $\frac{1}{\lambda}$, und erfüllt die Beschränktheitsbedingung für $C_\varphi = 1$. Der zugehörige Regularisierungsoperator ist gegeben durch

$$R_\alpha y = \sum_{n\in\mathbb{N}} \frac{\sigma_n}{\sigma_n^2 + \alpha}\,(y, u_n)_Y\, v_n,$$

die regularisierte Lösung $\varphi_\alpha(K^*K)K^*y$ lässt sich jedoch auch ohne Kenntnis einer Singulärwertzerlegung berechnen. Wir werden sie daher im nächsten Kapitel ausführlicher betrachten.

(iii) Die *Landweber-Regularisierung* entsteht durch die Wahl

$$\varphi_\alpha(\lambda) = \frac{1 - (1 - \omega\lambda)^{1/\alpha}}{\lambda}$$

für ein geeignetes $\omega > 0$. Auch hier existiert eine (intuitivere) Charakterisierung ohne Singulärwertzerlegung, weshalb wir die nähere Betrachtung auf ein folgendes Kapitel verschieben.

5.1 REGULARISIERUNG

Wir zeigen zuerst, dass für einen regularisierenden Filter $\{\varphi_\alpha\}_{\alpha>0}$ durch $R_\alpha := \varphi_\alpha(K^*K)K^*$ tatsächlich eine Regularisierung $\{R_\alpha\}_{\alpha>0}$ von $K^\dagger$ definiert wird. Dafür benötigen wir die folgenden drei fundamentalen Resultate, die wir im weiteren Verlauf immer wieder anwenden werden.

Lemma 5.3. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein regularisierender Filter. Dann gilt*

$$\|KR_\alpha\|_{\mathcal{L}(Y,Y)} \leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)|\sigma_n^2 \leq C_\varphi \qquad \text{für alle } \alpha > 0.$$

Beweis. Für alle $y \in Y$ und $\alpha > 0$ gilt (vergleiche (3.11))

$$\begin{aligned} KR_\alpha y = K\varphi_\alpha(K^*K)K^*y &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n \,(y,u_n)_Y \, Kv_n \\ &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n^2 \,(y,u_n)_Y \, u_n. \end{aligned} \tag{5.3}$$

Mit der Besselschen Ungleichung (2.2) folgt daher

$$\begin{aligned} \|KR_\alpha y\|_Y^2 = \sum_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)\sigma_n^2 \,(y,u_n)_Y\,|^2 &\leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)\sigma_n^2|^2 \sum_{n\in\mathbb{N}} |\,(y,u_n)_Y\,|^2 \\ &\leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)\sigma_n^2|^2 \|y\|_Y^2. \end{aligned}$$

Die zweite Ungleichung folgt nun aus $0 < \sigma_n^2 \leq \sigma_1^2 = \|K^*K\|_{\mathcal{L}(X,X)} = \kappa$ und der Eigenschaft (ii) von regularisierenden Filtern. □

Lemma 5.4. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein regularisierender Filter. Dann gilt*

$$\|R_\alpha\|_{\mathcal{L}(Y,X)} \leq \sqrt{C_\varphi} \sup_{\lambda\in(0,\kappa]} \sqrt{|\varphi_\alpha(\lambda)|} \qquad \text{für alle } \alpha > 0.$$

Insbesondere ist $R_\alpha : Y \longrightarrow X$ stetig für alle $\alpha > 0$.

Beweis. Für alle $y \in Y$ und $\alpha > 0$ folgt aus Lemma 5.3 wegen $\sigma_n v_n = K^*u_n$

$$\begin{aligned} \|R_\alpha y\|_X^2 = (R_\alpha y, R_\alpha y)_X &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n \,(y,u_n)_Y \,(R_\alpha y, v_n)_X \\ &= \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2) \,(y,u_n)_Y \,(KR_\alpha y, u_n)_Y \\ &\leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)| \left(KR_\alpha y, \textstyle\sum_{n\in\mathbb{N}} (y,u_n)_Y \, u_n\right)_Y \\ &\leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)| \; \|KR_\alpha y\|_X \|P_{\overline{\mathcal{R}(K^*)}} y\|_Y \\ &\leq \sup_{n\in\mathbb{N}} |\varphi_\alpha(\sigma_n^2)| \; C_\varphi \|y\|_Y^2. \end{aligned}$$

Supremum über alle $y \in Y$ sowie die Beschränktheit von φ_α ergeben nun die Behauptung. □

Das dritte „Fundamentallemma der spektralen Regularisierung“ gibt schließlich eine spektrale Darstellung des Verfahrensfehlers.

Lemma 5.5. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein regularisierender Filter und $y \in \mathcal{D}(K^\dagger)$ sowie $x^\dagger := K^\dagger y$. Dann gilt*

$$K^\dagger y - R_\alpha y = \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2) \left(x^\dagger, v_n\right)_X v_n$$

für $r_\alpha(\lambda) := 1 - \lambda\varphi_\alpha(\lambda)$ mit

$$\begin{aligned} &\lim_{\alpha\to 0} r_\alpha(\lambda) = 0 && \text{für alle } \lambda \in (0,\kappa], \\ &|r_\alpha(\lambda)| \le 1 + C_\varphi && \text{für alle } \lambda \in (0,\kappa] \text{ und } \alpha > 0. \end{aligned}$$

Beweis. Wegen $K^*Kx^\dagger = K^*y$ nach Folgerung 3.6 können wir schreiben

$$R_\alpha y = \varphi_\alpha(K^*K)K^*y = \varphi_\alpha(K^*K)K^*Kx^\dagger,$$

und mit der Definition von r_α und (3.16) für $x^\dagger \in \mathcal{N}(K)^\perp$ folgt sofort

$$K^\dagger y - R_\alpha y = (\mathrm{Id} - \varphi_\alpha(K^*K)K^*K)x^\dagger = r_\alpha(K^*K)x^\dagger = \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2)\left(x^\dagger, v_n\right)_X v_n.$$

Die restlichen Behauptungen folgen direkt aus den entsprechenden Eigenschaften von regularisierenden Filtern. □

Wir haben nun alles zur Hand, um die punktweise Konvergenz und damit die Regularisierungseigenschaft von $\{R_\alpha\}_{\alpha>0}$ zu zeigen.

Satz 5.6. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein regularisierender Filter. Dann gilt*

$$\lim_{\alpha\to 0} R_\alpha y = K^\dagger y \qquad \text{für alle } y \in \mathcal{D}(K^\dagger),$$

d. h. $\{R_\alpha\}_{\alpha>0}$ ist eine Regularisierung.

Ist $K^\dagger$ nicht stetig, gilt dagegen $\lim_{\alpha\to 0} \|R_\alpha y\|_X = \infty$ für alle $y \notin \mathcal{D}(K^\dagger)$.

Beweis. Sei $y \in \mathcal{D}(K^\dagger)$ und $x^\dagger = K^\dagger y$. Aus Lemma 5.5 folgt dann

$$\|K^\dagger y - R_\alpha y\|_X^2 = \sum_{n\in\mathbb{N}} |r_\alpha(\sigma_n^2)|^2 \left|\left(x^\dagger, v_n\right)_X\right|^2.$$

Um zu zeigen, dass die rechte Seite für $\alpha \to 0$ beliebig klein wird, spalten wir nun diese Reihe auf in einen endlichen Teil, für den wir die Konvergenz von r_α und die Beschränktheit der Skalarprodukte nutzen können, und ein Restglied, für das wir umgekehrt argumentieren.

Sei dafür $\varepsilon > 0$ beliebig. Dann existiert zuerst wegen der Besselschen Ungleichung ein $N \in \mathbb{N}$ mit

$$\sum_{n=N+1}^{\infty} \left| \left(x^\dagger, v_n \right)_X \right|^2 < \frac{\varepsilon^2}{2(1+C_\varphi)^2}.$$

Weiterhin existiert wegen der punktweisen Konvergenz von $\{r_\alpha\}_{\alpha>0}$ – die daher gleichmäßig ist auf der endlichen Menge $\{\sigma_1^2, \dots, \sigma_N^2\}$ – ein $\alpha_0 > 0$ mit

$$|r_\alpha(\sigma_n^2)|^2 < \frac{\varepsilon^2}{2\|x^\dagger\|_X^2} \qquad \text{für alle } n \le N \text{ und } \alpha < \alpha_0.$$

Für alle $\alpha < \alpha_0$ gilt daher

$$\begin{aligned} \|K^\dagger y - R_\alpha y\|_X^2 &= \sum_{n=1}^{N} |r_\alpha(\sigma_n^2)|^2 \left| \left(x^\dagger, v_n \right)_X \right|^2 + \sum_{n=N+1}^{\infty} |r_\alpha(\sigma_n^2)|^2 \left| \left(x^\dagger, v_n \right)_X \right|^2 \\ &\le \frac{\varepsilon^2}{2\|x^\dagger\|_X^2} \sum_{n=1}^{N} \left| \left(x^\dagger, v_n \right)_X \right|^2 + (1+C_\varphi)^2 \frac{\varepsilon^2}{2(1+C_\varphi)^2} \\ &\le \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2, \end{aligned}$$

d. h. $\|K^\dagger y - R_\alpha y\|_X \to 0$ für $\alpha \to 0$. Zusammen mit der Stetigkeit von R_α für $\alpha > 0$ aus Lemma 5.4 folgt, dass $\{R_\alpha\}_{\alpha>0}$ nach Definition 4.1 eine Regularisierung ist.

Die Unbeschränktheit für $y \notin \mathcal{D}(K^\dagger)$ folgt schließlich aus Satz 4.3 und Lemma 5.3. □

Damit definieren nach Beispiel 5.2 sowohl die abgeschnittene Singulärwertzerlegung, die Tikhonov-Regularisierung, als auch (nach Nachweis der Filter-Eigenschaften) die Landweber-Regularisierung für *beliebige* kompakte Operatoren eine Regularisierung.

5.2 PARAMETERWAHL UND KONVERGENZRATEN

Wir untersuchen nun, welche Parameterwahlstrategien α für einen gegebenen Filter φ_α zu einem konvergenten (und ordnungsoptimalen) Regularisierungsverfahren (R_α, α) führen. Der Übersichtlichkeit halber werden wir in Folge stets $x^\dagger := K^\dagger y$ und $x_\alpha := R_\alpha y$ für $y \in \mathcal{D}(K^\dagger)$ sowie $x_\alpha^\delta := R_\alpha y^\delta$ für $y^\delta \in B_\delta(y)$ schreiben.

A PRIORI-STRATEGIEN

Nach Satz 4.6 liefert jede a priori-Strategie mit $\alpha(\delta) \to 0$ und $\delta \|R_\alpha\|_{\mathcal{L}(Y,X)} \to 0$ für $\delta \to 0$ ein Regularisierungsverfahren (R_α, α). Zusammen mit Lemma 5.4 erhält man daraus eine Bedingung für φ_α und damit für α. Wir betrachten das an einem Beispiel.

Beispiel 5.7 (abgeschnittene Singulärwertzerlegung). Sei $K \in \mathcal{K}(X,Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$. Dann gilt für φ_α wie in (5.1)

$$\|R_\alpha\|_{\mathcal{L}(Y,X)} \leq \sqrt{C_\varphi} \sup_{n\in\mathbb{N}} \sqrt{|\varphi_\alpha(\sigma_n^2)|} = \frac{1}{\sqrt{\alpha}}.$$

Dies liefert eine Bedingung für den minimalen Singulärwert, den wir für ein gegebenes $\delta > 0$ in (5.2) berücksichtigen dürfen: Wähle für $\delta > 0$ ein $n(\delta)$ mit

$$n(\delta) \to \infty, \qquad \frac{\delta}{\sigma_{n(\delta)}} \to 0 \qquad \text{für } \delta \to 0.$$

Dann ergibt die abgeschnittene Singulärwertzerlegung zusammen mit der Parameterwahlstrategie $\alpha(\delta) := \sigma_{n(\delta)}^2$ ein Regularisierungsverfahren.

Dies gilt insbesondere für die Wahl $\alpha(\delta) := \sigma_{n(\delta)}^2 \geq \delta > \sigma_{n(\delta)+1}^2$, für die gilt

$$x_{\alpha(\delta)}^\delta = \sum_{\sigma_n \geq \sqrt{\delta}} \frac{1}{\sigma_n} \left(y^\delta, u_n\right)_Y v_n \to \sum_{n\in\mathbb{N}} \frac{1}{\sigma_n} (y, u_n)_Y v_n = x^\dagger \qquad \text{für } \delta \to 0.$$

Wir betrachten nun Konvergenzraten unter der Quellbedingung $x^\dagger \in X_{\nu,\rho}$ für $\nu, \rho > 0$. Wir gehen dafür wie im Beweis von Satz 5.6 vor und zeigen zunächst, dass

$$\omega_\nu(\alpha) := \sup_{\lambda\in(0,\kappa]} \lambda^{\nu/2} |r_\alpha(\lambda)|$$

eine obere Schranke für den Verfahrensfehler darstellt.

Lemma 5.8. *Sei $y \in \mathcal{D}(K^\dagger)$ und $x^\dagger \in X_{\nu,\rho}$. Dann gilt für alle $\alpha > 0$*

$$\|x_\alpha - x^\dagger\|_X \leq \omega_\nu(\alpha)\rho, \tag{5.4}$$

$$\|Kx_\alpha - Kx^\dagger\|_Y \leq \omega_{\nu+1}(\alpha)\rho. \tag{5.5}$$

Beweis. Für $x^\dagger \in X_{\nu,\rho}$ existiert ein $w \in X$ mit $x^\dagger = |K|^\nu w = (K^*K)^{\nu/2} w$ und $\|w\|_X \leq \rho$. Aus Lemma 5.5 folgt nun

$$\begin{aligned} x^\dagger - x_\alpha &= r_\alpha(K^*K)x^\dagger = r_\alpha(K^*K)(K^*K)^{\nu/2} w \\ &= \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2)\sigma_n^\nu \, (w, v_n)_X \, v_n \end{aligned}$$

und damit

$$\begin{aligned} \|x_\alpha - x^\dagger\|_X^2 &= \sum_{n\in\mathbb{N}} |r_\alpha(\sigma_n^2)|^2 \sigma_n^{2\nu} | \, (w, v_n)_X \, |^2 \\ &\leq \omega_\nu(\alpha)^2 \sum_{n\in\mathbb{N}} | \, (w, v_n)_X \, |^2 \leq \omega_\nu(\alpha)^2 \|w\|_X^2 \leq \omega_\nu(\alpha)^2 \rho^2. \end{aligned}$$

Nach Lemma 3.13 (iii) gilt weiterhin

$$\|Kx_\alpha - Kx^\dagger\|_Y = \|K(x_\alpha - x^\dagger)\|_Y = \||K|(x_\alpha - x^\dagger)\|_X,$$

woraus analog mit

$$\begin{aligned} |K|(x^\dagger - x_\alpha) &= (K^*K)^{1/2} r_\alpha(K^*K)(K^*K)^{\nu/2} w \\ &= \sum_{n\in\mathbb{N}} \sigma_n r_\alpha(\sigma_n^2)\sigma_n^\nu \,(w, v_n)_X \, v_n \end{aligned}$$

und $|r_\alpha(\sigma_n^2)\sigma_n^{\nu+1}|^2 \le \omega_{\nu+1}(\alpha)^2$ die zweite Abschätzung folgt. □

Damit haben wir jetzt alles zusammen, um Konvergenzraten zu zeigen.

Satz 5.9. *Sei $y \in \mathcal{D}(K^\dagger)$ und $x^\dagger = K^\dagger y \in X_{\nu,\rho}$. Sei $\alpha(\delta)$ eine Parameterwahlstrategie mit*

$$c\left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}} \le \alpha(\delta) \le C\left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}} \qquad \text{für } C > c > 0. \tag{5.6}$$

Erfüllt der Filter $\{\varphi_\alpha\}_{\alpha>0}$ für $\nu > 0$ die Bedingungen

$$\sup_{\lambda\in(0,\kappa]} |\varphi_\alpha(\lambda)| \le C_\varphi \alpha^{-1}, \tag{5.7}$$

$$\omega_\nu(\alpha) \le C_\nu \alpha^{\nu/2}, \tag{5.8}$$

für eine Konstante $C_\nu > 0$, so ist (R_α, α) für dieses ν ein ordnungsoptimales Regularisierungsverfahren.

Beweis. Nach Satz 4.14 genügt es, die Ordnungsoptimalität zu zeigen. Wir verwenden wieder die Zerlegung (4.2) in Datenfehler und Verfahrensfehler: Seien $\delta > 0$ und $y^\delta \in B_\delta(y)$ gegeben. Dann gilt

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\|_X \le \delta\|R_{\alpha(\delta)}\|_{\mathcal{L}(Y,X)} + \|x_{\alpha(\delta)} - x^\dagger\|_X.$$

Nach Lemma 5.4 und Annahme (5.7) ist nun

$$\|R_{\alpha(\delta)}\|_{\mathcal{L}(Y,X)} \le \sqrt{C_\varphi}\sqrt{C_\varphi \alpha(\delta)^{-1}} \le C_\varphi \alpha(\delta)^{-1/2}.$$

Analog folgt aus Lemma 5.8 und Annahme (5.8)

$$\|x_{\alpha(\delta)} - x^\dagger\|_X \le \omega_\nu(\alpha(\delta))\rho \le C_\nu \alpha(\delta)^{\nu/2}\rho.$$

Zusammen mit der Parameterwahlstrategie (5.6) erhalten wir daraus

$$\begin{aligned} \|x^\delta_{\alpha(\delta)} - x^\dagger\|_X &\le C_\varphi \alpha(\delta)^{-1/2}\delta + C_\nu \alpha(\delta)^{\nu/2}\rho \\ &\le C_\varphi c^{-1/2}\delta^{-\frac{1}{\nu+1}}\rho^{\frac{1}{\nu+1}}\delta + C_\nu C^{\nu/2}\delta^{\frac{\nu}{\nu+1}}\rho^{-\frac{\nu}{\nu+1}}\rho \\ &= (C_\varphi c^{-1/2} + C_\nu C^{\nu/2})\delta^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}, \end{aligned} \tag{5.9}$$

und damit die Ordnungsoptimalität. Da $\alpha_\tau(\delta, y^\delta) := \alpha(\tau\delta, y^\delta)$ für alle $\tau > 0$ ebenfalls die Form (5.6) (mit von τ abhängenden Konstanten c, C) haben, folgt aus Satz 4.14 auch die Konvergenz für alle $y \in \mathcal{D}(K^\dagger)$. □

Um für einen konkreten Filter die Ordnungsoptimalität für ein $\nu > 0$ zu zeigen, genügt es also, für dieses ν die Bedingung (5.8) (sowie für φ_α die Bedingung (5.7)) nachzuweisen. Das maximale $\nu_0 > 0$, so dass für alle $\nu \in (0, \nu_0]$ die Bedingung (5.8) gilt, heißt *Qualifikation* des Filters.

Beispiel 5.10 (abgeschnittene Singulärwertzerlegung). Aus (5.1) folgt

$$\sup_{\lambda\in(0,\kappa]} |\varphi_\alpha(\lambda)| \le \alpha^{-1},$$

womit wegen $C_\varphi = 1$ für diesen Filter die Bedingung (5.7) erfüllt ist.

Weiterhin gilt für alle $\nu > 0$ und $\lambda \in (0, \kappa]$

$$\lambda^{\nu/2}|r_\alpha(\lambda)| = \lambda^{\nu/2}|1 - \lambda\varphi_\alpha(\lambda)| = \begin{cases} 0 & \text{falls } \lambda \ge \alpha, \\ \lambda^{\nu/2} & \text{falls } \lambda < \alpha. \end{cases}$$

Für alle $\alpha \in (0, \kappa]$ ist also

$$\omega_\nu(\alpha) = \sup_{\lambda\in(0,\kappa]} \lambda^{\nu/2}|r_\alpha(\lambda)| \le \max\{0, \alpha^{\nu/2}\} = \alpha^{\nu/2},$$

und die Bedingung (5.8) ist damit für alle $\nu > 0$ mit $C_\nu = 1$ erfüllt. (Für $\alpha > \kappa$ werden *alle* Singulärwerte abgeschnitten, d. h. $R_\alpha = 0$.) Die abgeschnittene Singulärwertzerlegung ist daher für alle $\nu > 0$ ordnungsoptimal; man sagt, sie besitzt *unendliche Qualifikation*.

A POSTERIORI-STRATEGIEN

Wir betrachten wieder das Diskrepanzprinzip: Für $\tau > 1$ sei $\alpha(\delta, y^\delta)$ so bestimmt, dass gilt

$$\|Kx^\delta_{\alpha(\delta,y^\delta)} - y^\delta\|_Y \le \tau\delta < \|Kx^\delta_\alpha - y^\delta\|_Y \qquad \text{für alle } \alpha > \alpha(\delta, y^\delta). \tag{5.10}$$

Wieder nehmen wir an, dass $\mathcal{R}(K)$ dicht in Y ist. Ist weiterhin $\alpha \mapsto \varphi_\alpha(\lambda)$ stetig für alle $\lambda \in (0, \kappa]$, so kann man daraus die Stetigkeit von $\alpha \mapsto \|Kx^\delta_\alpha - y^\delta\|_Y$ folgern und damit analog zu Satz 4.7 mit Hilfe von Lemma 5.3 zeigen, dass so ein $\alpha(\delta, y^\delta)$ stets existiert. Um zu zeigen, dass das Diskrepanzprinzip zu einem Regularisierungsverfahren führt, verwenden wir wieder Satz 4.14, wobei wir hier das Diskrepanzprinzip als Wahlregel $\alpha_\tau = \alpha(\tau\delta, y^\delta)$ auffassen müssen.

Satz 5.11. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein Filter mit Qualifikation $\nu_0 > 0$ (d. h. erfüllt (5.7) und (5.8) für alle $\nu \in (0, \nu_0]$) sowie*

$$\tau > \sup_{\alpha>0,\, \lambda\in(0,\kappa]} |r_\alpha(\lambda)| =: C_r. \tag{5.11}$$

Dann definiert das Diskrepanzprinzip für alle $\nu \in (0, \nu_0 - 1]$ ein ordnungsoptimales Regularisierungsverfahren (R_α, α_τ).

Beweis. Zuerst halten wir fest, dass wegen $|r_\alpha(\lambda)| \le 1 + C_\varphi$ für alle $\alpha > 0$ und $\lambda \in (0, \kappa]$ stets ein $\tau > 1$ mit (5.11) existiert.

Seien nun $y \in \mathcal{R}(K)$, $x^\dagger = K^\dagger y \in X_{\nu,\rho}$ für ein $\nu \in (0, \nu_0 - 1]$ und $\rho > 0$, und $y^\delta \in B_\delta(y)$. Wir verwenden für $x_\alpha^\delta := x^\delta_{\alpha(\delta,y^\delta)}$ und $x_\alpha := x_{\alpha(\delta,y^\delta)}$ wieder die Zerlegung

$$\|x_\alpha^\delta - x^\dagger\|_X \le \|x_\alpha - x^\dagger\|_X + \|x_\alpha - x_\alpha^\delta\|_X \tag{5.12}$$

und schätzen die Summanden separat ab.

Für den ersten Term verwenden wir wieder die Darstellung des Verfahrensfehlers aus Lemma 5.5 sowie die Quelldarstellung $x^\dagger = |K|^\nu w$ und erhalten

$$\begin{aligned} x^\dagger - x_\alpha &= \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2)\sigma_n^\nu \,(w, v_n)_X \, v_n \\ &= \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2)\, (w, v_n)_X \, |K|^\nu v_n \\ &= |K|^\nu \sum_{n\in\mathbb{N}} r_\alpha(\sigma_n^2)\, (w, v_n)_X \, v_n =: |K|^\nu \xi. \end{aligned}$$

Die Interpolationsungleichung (3.18) mit $r = \nu$, $s = \nu + 1$ liefert dann

$$\|x_\alpha - x^\dagger\|_X = \||K|^\nu \xi\|_X \le \||K|^{\nu+1}\xi\|_X^{\frac{\nu}{\nu+1}} \, \|\xi\|_X^{\frac{1}{\nu+1}}.$$

Für den zweiten Faktor erhalten wir aus der Definition von ξ zusammen mit der Besselschen Ungleichung, der Beschränktheit von r_α, und der Quellbedingung

$$\|\xi\|_X^2 = \sum_{n\in\mathbb{N}} |r_\alpha(\sigma_n^2)|^2 |\, (w, v_n)_X \,|^2 \le C_r^2 \|w\|_X^2 \le C_r^2 \rho^2.$$

Für den ersten Faktor verwenden wir Lemma 3.13 (i), (iii), $Kx^\dagger = y$ wegen $y \in \mathcal{R}(K)$, und die produktive Null:

$$\begin{aligned} \||K|^{\nu+1}\xi\|_X &= \||K|(|K|^\nu \xi)\|_X = \||K|(x_\alpha - x^\dagger)\|_X = \|K(x_\alpha - x^\dagger)\|_Y = \|Kx_\alpha - y\|_Y \\ &\le \|Kx_\alpha^\delta - y^\delta\|_Y + \|y - y^\delta - K(x_\alpha - x_\alpha^\delta)\|_Y. \end{aligned}$$

Wieder schätzen wir separat ab: wegen der Wahl von $\alpha(\delta, y^\delta)$ nach dem Diskrepanzprinzip ist zuerst $\|Kx_\alpha^\delta - y^\delta\|_Y \le \tau\delta$. Für den zweiten Summanden gehen wir wie folgt vor: Es gilt

$$y - Kx_\alpha = y - KR_\alpha y = (\mathrm{Id} - K\varphi_\alpha(K^*K)K^*)y$$

und analog für $y^\delta - Kx_\alpha^\delta$. Also ist

$$\begin{aligned} \|y - y^\delta - K(x_\alpha - x_\alpha^\delta)\|_Y^2 &= \|(\mathrm{Id} - K\varphi_\alpha(K^*K)K^*)(y - y^\delta)\|_Y^2 \\ &= \sum_{n\in\mathbb{N}} |r_\alpha(\sigma_n^2)|^2 \left|\left(y - y^\delta, u_n\right)_Y\right|^2 \\ &\le C_r^2\delta^2, \end{aligned} \tag{5.13}$$

wobei in der zweiten Gleichung verwendet wurde, dass gilt (vergleiche (5.3))

$$K\varphi_\alpha(K^*K)K^*(y - y^\delta) = \sum_{n\in\mathbb{N}} \varphi_\alpha(\sigma_n^2)\sigma_n^2 \left(y - y^\delta, u_n\right)_Y u_n.$$

Insgesamt erhalten wir also für den ersten Term in (5.12)

$$\|x_\alpha - x^\dagger\|_X \le (\tau + C_r)^{\frac{\nu}{\nu+1}} \delta^{\frac{\nu}{\nu+1}} C_r^{\frac{1}{\nu+1}} \rho^{\frac{1}{\nu+1}} =: C_1 \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}.$$

Für den zweiten Term in (5.12) verwenden wir Lemma 5.4 und Annahme (5.7) und erhalten

$$\begin{aligned} \|x_\alpha^\delta - x_\alpha\|_X \le \|R_\alpha\|_{\mathcal{L}(Y,X)}\delta &\le \sqrt{C_\varphi} \sup_{\lambda\in(0,\kappa]} \sqrt{|\varphi_\alpha(\lambda)|}\delta \\ &\le C_\varphi \alpha(\delta, y^\delta)^{-1/2}\delta. \end{aligned} \tag{5.14}$$

Um auf die gewünschte Ordnung zu kommen, brauchen wir nun eine Abschätzung für $\alpha(\delta, y^\delta)$. Aufgrund der Wahl nach dem Diskrepanzprinzip gilt insbesondere

$$\tau\delta < \|Kx_{2\alpha}^\delta - y^\delta\|_Y \le \|Kx_{2\alpha} - y\|_Y + \|y - y^\delta - K(x_{2\alpha} - x_{2\alpha}^\delta)\|_Y \tag{5.15}$$

(die Wahl $2\alpha > \alpha$ ist dabei willkürlich). Da die Abschätzung (5.13) gleichmäßig in $\alpha > 0$ ist, gilt für den letzten Term ebenfalls

$$\|y - y^\delta - K(x_{2\alpha} - x_{2\alpha}^\delta)\|_Y \le C_r\delta$$

und damit

$$\|Kx_{2\alpha} - y\|_Y > \tau\delta - \|y - y^\delta - K(x_{2\alpha} - x_{2\alpha}^\delta)\|_Y \ge (\tau - C_r)\delta.$$

In die andere Richtung schätzen wir mit Lemma 5.8 und Annahme (5.8) für $\nu + 1 \le \nu_0$ ab:

$$\|Kx_{2\alpha} - y\|_Y \le \omega_{\nu+1}(2\alpha(\delta, y^\delta))\rho \le C_{\nu+1}(2\alpha(\delta, y^\delta))^{\frac{\nu+1}{2}}\rho.$$

Zusammen mit der Annahme $\tau > C_r$ folgt daraus

$$\delta < (\tau - C_r)^{-1} C_{\nu+1} 2^{\frac{\nu+1}{2}} \alpha(\delta, y^\delta)^{\frac{\nu+1}{2}} \rho =: C_\tau \alpha(\delta, y^\delta)^{\frac{\nu+1}{2}} \rho,$$

d. h.

$$\alpha(\delta, y^\delta)^{-1/2} \leq C_\tau^{\frac{1}{\nu+1}} \delta^{-\frac{1}{\nu+1}} \rho^{\frac{1}{\nu+1}}. \tag{5.16}$$

Einsetzen in (5.14) ergibt dann

$$\|x_\alpha^\delta - x_\alpha\|_X \leq C_\varphi C_\tau^{\frac{1}{\nu+1}} \delta \delta^{-\frac{1}{\nu+1}} \rho^{\frac{1}{\nu+1}} =: C_2 \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}.$$

Insgesamt erhalten wir also

$$\|x_\alpha^\delta - x^\dagger\|_X \leq (C_1 + C_2) \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}$$

und damit die Ordnungsoptimalität. Aus Satz 4.14 für $\nu = \nu_0 - 1$ und $\tau_0 = C_r$ folgt damit auch, dass R_α zusammen mit dem Diskrepanzprinzip als Parameterwahlstrategie $\alpha_\tau = \alpha(\tau\delta, y^\delta)$ für alle $\tau > C_r$ ein Regularisierungsverfahren definiert. □

Beispiel 5.12 (abgeschnittene Singulärwertzerlegung). Hier gilt

$$|r_\alpha(\lambda)| = \begin{cases} 1 - \lambda \frac{1}{\lambda} = 0 & \lambda \geq \alpha \\ 1 & \lambda < \alpha \end{cases}$$

und damit $C_r = 1$. Wegen der unendlichen Qualifikation ist daher die abgeschnittene Singulärwertzerlegung zusammen mit dem Diskrepanzprinzip für beliebige $\tau > 1$ und $\nu > 0$ ein ordnungsoptimales Regularisierungsverfahren.

Hat ein Filter jedoch nur endliche Qualifikation, führt das Diskrepanzprinzip für $\nu > \nu_0 - 1$ nicht zu einem ordnungsoptimalen Regularisierungsverfahren. Es gibt jedoch verbesserte Diskrepanzprinzipien, die das Residuum in angepassten Normen messen und dadurch zu ordnungsoptimalen Verfahren auch für $\nu \in (\nu_0 - 1, \nu_0]$ führen; siehe z. B. [Engl, Hanke & Neubauer 1996, Kapitel 4.4].

HEURISTISCHE STRATEGIEN

Wir behandeln als Beispiel die Hanke–Raus-Regel: Betrachte für $y^\delta \in Y$ die Funktion

$$\Psi : (0, \kappa] \to \mathbb{R}, \qquad \Psi(\alpha) = \frac{\|K x_\alpha^\delta - y^\delta\|_Y}{\sqrt{\alpha}},$$

und wähle

$$\alpha(y^\delta) \in \arg\min_{\alpha \in (0,\kappa]} \Psi(\alpha). \tag{5.17}$$

Wir nehmen in Folge an, dass $y \in \mathcal{R}(K)$ sowie $\|y\|_Y > \delta$ gelten. Wir zeigen zunächst eine bedingte Fehlerabschätzung.

Satz 5.13. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein Filter mit Qualifikation $\nu_0 > 0$, d. h. es gelte (5.7) sowie (5.8) für alle $\nu \in (0, \nu_0]$. Es existiere weiterhin ein Minimierer $\alpha^* := \alpha(y^\delta) \in (0, \kappa]$ von Ψ mit*

$$\delta^* := \|Kx^\delta_{\alpha^*} - y^\delta\|_Y > 0. \tag{5.18}$$

Dann existiert ein $c > 0$ so dass für alle $x^\dagger \in X_{\nu,\rho}$ mit $\nu \in (0, \nu_0 - 1]$ und $\rho \geq 0$ gilt

$$\|x^\delta_{\alpha^*} - x^\dagger\|_X \leq c\left(1 + \frac{\delta}{\delta^*}\right)\max\{\delta, \delta^*\}^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}.$$

Beweis. Wir verwenden wieder die Zerlegung

$$\|x^\delta_{\alpha^*} - x^\dagger\|_X \leq \|x_{\alpha^*} - x^\dagger\|_X + \|x^\delta_{\alpha^*} - x_{\alpha^*}\|_X.$$

Für den ersten Term erhalten wir wie im Beweis von Satz 5.11 mit (5.18) anstelle des Diskrepanzprinzips

$$\|x_{\alpha^*} - x^\dagger\|_X \leq C_r^{\frac{1}{\nu+1}}(\delta^* + C_r\delta)^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}} \leq C_1\max\{\delta, \delta^*\}^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}. \tag{5.19}$$

für eine Konstante $C_1 > 0$.

Für den zweiten Term erhalten wir analog zu (5.14) mit Hilfe von (5.18) (in Form der produktiven $1 = \delta^*/\delta^*$)

$$\|x^\delta_{\alpha^*} - x_{\alpha^*}\|_X \leq C_\varphi\frac{1}{\sqrt{\alpha^*}}\delta = C_\varphi\frac{\delta}{\delta^*}\frac{\|Kx^\delta_{\alpha^*} - y^\delta\|_Y}{\sqrt{\alpha^*}} = C_\varphi\frac{\delta}{\delta^*}\Psi(\alpha^*).$$

Für den letzten Faktor verwenden wir nun die Wahlregel (5.17): Es gilt $\Psi(\alpha^*) \leq \Psi(\alpha)$ für alle $\alpha \in (0, \kappa]$. Die Idee ist nun, zum Vergleich das Diskrepanzprinzip heranzuziehen. Sei $\bar\alpha := \alpha(\delta, y^\delta)$ so gewählt, dass (5.10) für ein $\tau > C_r$ nach (5.11) erfüllt ist. Ist $\bar\alpha \leq \kappa$, so gilt wegen (5.16)

$$\Psi(\alpha^*) \leq \Psi(\bar\alpha) \leq (\tau\delta)(C_\tau^{\frac{1}{\nu+1}}\delta^{-\frac{1}{\nu+1}}\rho^{\frac{1}{\nu+1}}) = C_\tau^{\frac{1}{\nu+1}}\tau\delta^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}. \tag{5.20}$$

Ist dagegen $\bar\alpha > \kappa = \|K\|^2_{\mathcal{L}(X,Y)}$, so ist nach Annahme auch $\|Kx^\delta_\kappa - y^\delta\|_Y \leq \tau\delta$. Wegen

$$\delta < \|y\|_Y = \|Kx^\dagger\|_Y = \|K|K|^\nu w\|_X \leq \|K\|^{\nu+1}_{\mathcal{L}(X,Y)}\rho$$

können wir daher abschätzen

$$\Psi(\alpha^*) \leq \Psi(\kappa) \leq \tau\delta\|K\|^{-1}_{\mathcal{L}(X,Y)} < \tau\delta\left(\delta^{-\frac{1}{\nu+1}}\rho^{\frac{1}{\nu+1}}\right) = \tau\delta^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}. \tag{5.21}$$

In beiden Fällen erhalten wir also

$$\|x^\delta_{\alpha^*} - x_{\alpha^*}\|_X \leq C_2\frac{\delta}{\delta^*}\delta^{\frac{\nu}{\nu+1}}\rho^{\frac{1}{\nu+1}}$$

für eine Konstante $C_2 > 0$, und zusammen mit (5.19) folgt die gewünschte Abschätzung. □

Die Hanke–Raus-Regel wäre also ordnungsoptimal, falls $\delta^* \approx \delta$ gelten würde. Umgekehrt wird die Regel versagen für $\alpha^* = 0$ oder $\delta^* = 0$. In letzterem Fall wäre $y^\delta \in \mathcal{R}(K)$, und wegen der Unstetigkeit von $K^\dagger$ kann $\|K^\dagger y^\delta - K^\dagger y\|_Y$ beliebig groß sein. Diesen Fall müssen wir daher ausschließen, um weitere Aussagen treffen zu können. Wir nehmen zum Beispiel an, dass ein $\varepsilon > 0$ existiert, so dass

$$y^\delta \in \mathcal{N}_\varepsilon := \left\{ y + \eta \in Y : \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})\eta\|_Y \geq \varepsilon \|\eta\|_Y \right\} \tag{5.22}$$

gilt, wobei $P_{\overline{\mathcal{R}}}$ wieder die Projektion auf $\overline{\mathcal{R}(K)}$ bedeutet. Anschaulich bedeutet das, dass die gestörten Daten y^δ nicht beliebig nahe an $\overline{\mathcal{R}(K)}$ liegen können. Wenn wir uns auf solche Störungen beschränken, liefert die Hanke–Raus-Regel tatsächlich ein konvergentes Regularisierungsverfahren.

Satz 5.14. *Sei $\{\varphi_\alpha\}_{\alpha>0}$ ein Filter mit Qualifikation $\nu_0 > 0$, d. h. es gelte (5.7) sowie (5.8) für alle $\nu \in (0, \nu_0]$. Dann gilt für alle $x^\dagger \in X_{\nu,\rho}$ mit $\nu \in (0, \nu_0 - 1]$ und $\rho > 0$*

$$\lim_{\delta \to 0} \sup_{y^\delta \in B_\delta(Kx^\dagger) \cap \mathcal{N}_\varepsilon} \|x^\delta_{\alpha^*} - x^\dagger\|_X = 0.$$

Beweis. Sei $y \in \mathcal{R}(K)$ und $y^\delta \in \mathcal{N}_\varepsilon$ mit $\|y^\delta - y\|_Y = \delta$. Da $\mathrm{Id} - P_{\overline{\mathcal{R}}}$ als orthogonale Projektion Norm 1 hat, gilt für alle $\alpha > 0$

$$\begin{aligned} \|Kx^\delta_\alpha - y^\delta\|_Y &\geq \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})(Kx^\delta_\alpha - y^\delta)\|_Y = \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})y^\delta\|_Y \\ &= \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})(y^\delta - y)\|_Y \geq \varepsilon \|y^\delta - y\|_Y \\ &= \varepsilon\delta > 0. \end{aligned} \tag{5.23}$$

Also ist der Zähler von $\Psi(\alpha)$ nach unten beschränkt, und daher gilt $\Psi(\alpha) \to \infty$ für $\alpha \to 0$. Das Infimum über alle $(0, \kappa]$ muss also für $\alpha^* > 0$ angenommen werden. Insbesondere folgt aus (5.23) auch

$$\delta^* = \|Kx^\delta_{\alpha^*} - y^\delta\|_Y \geq \varepsilon\delta > 0.$$

Satz 5.13 liefert daher mit $\delta \leq \varepsilon^{-1}\delta^*$ die Abschätzung

$$\|x^\delta_{\alpha^*} - x^\dagger\|_X \leq C_\varepsilon (\delta^*)^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}$$

für eine Konstante $C_\varepsilon > 0$. Es genügt also zu zeigen, dass für $\delta \to 0$ auch $\delta^* \to 0$ geht. Dies folgt aber aus $\alpha^* \leq \kappa$ und (5.20) bzw. (5.21), denn für $\delta \to 0$ gilt

$$\delta^* = \|Kx^\delta_{\alpha^*} - y^\delta\|_Y = \sqrt{\alpha^*}\Psi(\alpha^*) \leq \sqrt{\kappa}\Psi(\alpha^*) \leq \sqrt{\kappa} \max\{1, C_\tau^{\frac{1}{\nu+1}}\} \tau \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}} \to 0. \qquad \square$$

Unter ähnlichen Annahmen kann man mit mehr Aufwand auch die Ordnungsoptimalität für die Hanke–Raus-Regel zeigen, ebenso für verwandte Regeln, die auf der Minimierung von geeigneten Funktionalen beruhen; siehe etwa [Kindermann 2011].

6 TIKHONOV-REGULARISIERUNG

Aufgrund ihrer zentralen Rolle in der Theorie und Praxis der inversen Probleme betrachten wir noch einmal ausführlicher die Tikhonov-Regularisierung, die definiert wird durch die Filterfunktion

$$\varphi_\alpha(\lambda) = \frac{1}{\lambda + \alpha}.$$

Wir kommen schnell zur Sache, denn wir sind gut vorbereitet. Wie in Beispiel 5.2 (ii) bereits bemerkt, ist φ_α stetig, konvergiert für $\alpha \to 0$ gegen $\frac{1}{\lambda}$, ist gleichmäßig beschränkt durch α^{-1} und erfüllt

$$\lambda\varphi_\alpha(\lambda) = \frac{\lambda}{\lambda + \alpha} < 1 =: C_\varphi \qquad \text{für alle } \alpha > 0.$$

Damit ist nach Satz 5.6 $R_\alpha = \varphi_\alpha(K^*K)K^*$ ein Regularisierungsoperator, der nach Lemma 5.4

$$\|R_\alpha\|_{\mathcal{L}(Y,X)} \le \frac{1}{\sqrt{\alpha}}$$

erfüllt und daher nach Satz 4.6 zusammen mit der a priori-Strategie $\alpha(\delta) = \delta$ ein konvergentes Regularisierungsverfahren ergibt.

Um Konvergenzraten zu zeigen, verwenden wir Satz 5.9 bzw. Satz 5.11. Zunächst ist wegen $\varphi_\alpha(\lambda) \le \alpha^{-1} = C_\varphi \alpha^{-1}$ für alle $\alpha > 0$ die Bedingung (5.7) erfüllt. Weiterhin ist

$$r_\alpha(\lambda) = 1 - \lambda\varphi_\alpha(\lambda) = \frac{\alpha}{\lambda + \alpha} \le 1 =: C_r \qquad \text{für alle } \alpha > 0, \lambda \in (0, \kappa].$$

Um die zweite Bedingung (5.8) nachzuweisen, müssen wir

$$\omega_\nu(\alpha) = \sup_{\lambda\in(0,\kappa]} \lambda^{\nu/2}|r_\alpha(\lambda)| = \sup_{\lambda\in(0,\kappa]} \frac{\lambda^{\nu/2}\alpha}{\lambda + \alpha} =: \sup_{\lambda\in(0,\kappa]} h_\alpha(\lambda)$$

durch $C_\nu \alpha^{\nu/2}$ für eine Konstante $C_\nu > 0$ abschätzen. Dafür betrachten wir $h_\alpha(\lambda)$ für festes $\alpha > 0$ als Funktion von λ und berechnen

$$h'_\alpha(\lambda) = \frac{\alpha\frac{\nu}{2}\lambda^{\nu/2-1}(\lambda + \alpha) - \alpha\lambda^{\nu/2}}{(\lambda + \alpha)^2} = \frac{\alpha\lambda^{\nu/2-1}}{(\lambda + \alpha)^2}\left(\frac{\nu}{2}\alpha + \left(\frac{\nu}{2} - 1\right)\lambda\right).$$

Für $\nu \geq 2$ ist daher $h_\alpha(\lambda)$ monoton wachsend in λ, und das Maximum über alle $\lambda \in (0, \kappa]$ wird für $\lambda^* := \kappa$ angenommen. In diesem Fall ist

$$\omega_\nu(\alpha) = h_\alpha(\kappa) = \frac{\alpha \kappa^{\nu/2}}{\kappa + \alpha} \leq \kappa^{\nu/2-1}\alpha.$$

Die rechte Seite hat also nur für $\nu = 2$ die gewünschte Ordnung.

Für $\nu \in (0, 2)$ können wir die Nullstelle von $h'_\alpha(\lambda)$ bestimmen als $\lambda^* := \frac{\alpha \frac{\nu}{2}}{1-\frac{\nu}{2}}$. Dort ist $h''_\alpha(\lambda^*) < 0$, also ist für alle $\alpha > 0$

$$\omega_\nu(\alpha) = h_\alpha(\lambda^*) = \frac{\alpha \left(\alpha \frac{\nu}{2} \left(1 - \frac{\nu}{2}\right)^{-1}\right)^{\nu/2}}{\alpha + \alpha \frac{\nu}{2} \left(1 - \frac{\nu}{2}\right)^{-1}} \leq \left(\frac{\nu}{2}\left(1 - \frac{\nu}{2}\right)^{-1}\right)^{\nu/2} \alpha^{\nu/2},$$

womit wir die gewünschte Ordnung erhalten.

Die Tikhonov-Regularisierung hat also mindestens (und wie wir zeigen werden, höchstens) Qualifikation $\nu_0 = 2$. Aus Satz 5.9 und Satz 5.11 folgt nun die Ordnungsoptimalität für a priori- und a posteriori-Strategien.

Folgerung 6.1. *Für alle $\nu \in (0, 2]$ ist die Tikhonov-Regularisierung zusammen mit der Parameterwahlstrategie*

$$c\left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}} \leq \alpha(\delta) \leq C\left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}} \qquad \text{für } C > c > 0$$

ein ordnungsoptimales Regularisierungsverfahren. Insbesondere gilt für $\alpha \sim \delta^{2/3}$

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\|_X \leq c\delta^{\frac{2}{3}} \qquad \text{für alle } x^\dagger \in \mathcal{R}(K^*K) \text{ und } y^\delta \in B_\delta(Kx^\dagger).$$

Folgerung 6.2. *Für alle $\nu \in (0, 1]$ und $\tau > 1$ ist die Tikhonov-Regularisierung zusammen mit der Parameterwahlstrategie*

$$\|Kx^\delta_{\alpha(\delta, y^\delta)} - y^\delta\|_Y \leq \tau\delta < \|Kx^\delta_\alpha - y^\delta\|_Y \qquad \text{für alle } \alpha > \alpha(\delta, y^\delta)$$

ein ordnungsoptimales Regularisierungsverfahren. Insbesondere gilt

$$\|x^\delta_{\alpha(\delta, y^\delta)} - x^\dagger\|_X \leq c\delta^{\frac{1}{2}} \qquad \text{für alle } x^\dagger \in \mathcal{R}(K^*) \text{ und } y^\delta \in B_\delta(Kx^\dagger).$$

Tatsächlich kann die Qualifikation nicht höher als 2 sein; die Tikhonov-Regularisierung *sättigt* im Gegensatz zur abgeschnittenen Singulärwertzerlegung. Um dies zu zeigen, leiten wir die in Beispiel 5.2 (ii) versprochene alternative Charakterisierung her.

Lemma 6.3. *Für $y \in Y$ und $\alpha > 0$ ist $x_\alpha := R_\alpha y$ eindeutig bestimmt als Lösung der Gleichung*

$$(K^*K + \alpha \operatorname{Id})x = K^*y. \tag{6.1}$$

Beweis. Wir verwenden das singuläre System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ von K und erhalten

$$\alpha x_\alpha = \sum_{n\in\mathbb{N}} \alpha \frac{\sigma_n}{\sigma_n^2 + \alpha} (y, u_n)_Y v_n$$

sowie

$$\begin{aligned} K^*Kx_\alpha &= \sum_{n\in\mathbb{N}} \frac{\sigma_n}{\sigma_n^2 + \alpha} (y, u_n)_Y K^*Kv_n \\ &= \sum_{n\in\mathbb{N}} \sigma_n^2 \frac{\sigma_n}{\sigma_n^2 + \alpha} (y, u_n)_Y v_n. \end{aligned}$$

Daraus folgt

$$(K^*K + \alpha \operatorname{Id})x_\alpha = \sum_{n\in\mathbb{N}} \sigma_n (y, u_n)_Y v_n = K^*y.$$

Sei umgekehrt $x \in X$ eine Lösung von (6.1). Einsetzen der Darstellung

$$x = \sum_{n\in\mathbb{N}} (x, v_n)_X v_n + P_{\mathcal{N}}x \tag{6.2}$$

in (6.1) ergibt dann

$$\sum_{n\in\mathbb{N}} (\sigma_n^2 + \alpha) (x, v_n)_X v_n + \alpha P_{\mathcal{N}}x = (K^*K + \alpha \operatorname{Id})x = K^*y = \sum_{n\in\mathbb{N}} \sigma_n (y, u_n)_Y v_n.$$

Da $\{v_n\}_{n\in\mathbb{N}}$ eine Orthonormalbasis von $\overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$ ist, muss $P_{\mathcal{N}}x = 0$ sein. Durch Koeffizientenvergleich folgt dann

$$(x, v_n)_X = \frac{\sigma_n}{\sigma_n^2 + \alpha} (y, u_n)_Y \qquad \text{für alle } n \in \mathbb{N}.$$

Einsetzen in (6.2) ergibt wiederum

$$x = \sum_{n\in\mathbb{N}} (x, v_n)_X v_n = \sum_{n\in\mathbb{N}} \frac{\sigma_n}{\sigma_n^2 + \alpha} (y, u_n)_Y v_n = x_\alpha,$$

d. h. x_α ist die eindeutige Lösung von (6.1). □

Der praktische Wert der Darstellung (6.1) kann nicht genug betont werden: Anstelle der Singulärwertzerlegung muss lediglich die Lösung einer *korrekt gestellten* linearen Gleichung (für einen selbstadjungierten und positiv definiten Operator) berechnet werden, wofür Standardverfahren eingesetzt werden können.

Wir zeigen nun, dass im Allgemeinen keine a priori-Strategie existieren kann, für die der Fehler $\|x^\delta_{\alpha(\delta)} - x^\dagger\|_X$ schneller als $\delta^{2/3}$ gegen Null geht.

Satz 6.4. *Sei $K \in \mathcal{K}(X,Y)$ mit unendlichdimensionalem Bild und $y \in \mathcal{R}(K)$. Existiert eine a priori Parameterwahlstrategie α mit $\lim_{\delta\to 0} \alpha(\delta) = 0$ und*

$$\lim_{\delta\to 0} \sup_{y^\delta \in B_\delta(y)} \|x^\delta_{\alpha(\delta)} - x^\dagger\|_X \delta^{-\frac{2}{3}} = 0, \tag{6.3}$$

so ist $x^\dagger = 0$.

Beweis. Angenommen, $x^\dagger \neq 0$. Wir zeigen zuerst, dass unter diesen Voraussetzungen $\alpha(\delta)\delta^{-2/3} \to 0$ gehen muss. Dafür verwenden wir die Charakterisierung (6.1) für x^δ_α und y^δ, indem wir schreiben

$$(K^*K + \alpha(\delta)\,\mathrm{Id})\left(x^\delta_{\alpha(\delta)} - x^\dagger\right) = K^*y^\delta - K^*y - \alpha(\delta)x^\dagger.$$

Daraus folgt mit $\kappa = \|K^*K\|_{\mathcal{L}(X,X)} = \|K^*\|^2_{\mathcal{L}(Y,X)}$

$$|\alpha(\delta)|\|x^\dagger\|_X \leq \sqrt{\kappa}\delta + (\alpha(\delta)+\kappa)\|x^\delta_{\alpha(\delta)} - x^\dagger\|_X.$$

Multiplizieren mit $\delta^{-2/3}$ und Anwenden der Annahmen (6.3) und $x^\dagger \neq 0$ liefert nun

$$|\alpha(\delta)|\delta^{-2/3} \leq \|x^\dagger\|_X^{-1}\left(\sqrt{\kappa}\delta^{\frac{1}{3}} + (\alpha(\delta)+\kappa)\|x^\delta_{\alpha(\delta)} - x^\dagger\|_X\delta^{-\frac{2}{3}}\right) \to 0.$$

Wir konstruieren nun einen Widerspruch. Sei $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ ein singuläres System von K und definiere

$$\delta_n := \sigma_n^3 \quad \text{und} \quad y_n := y + \delta_n u_n, \qquad n \in \mathbb{N},$$

so dass gilt $\|y_n - y\|_Y = \delta_n \to 0$ für $n \to \infty$. Sei weiter $\alpha_n := \alpha(\delta_n)$. Dann ist

$$\begin{aligned} x^{\delta_n}_{\alpha_n} - x^\dagger &= (x^{\delta_n}_{\alpha_n} - x_{\alpha_n}) + (x_{\alpha_n} - x^\dagger) \\ &= R_\alpha(y_n - y) + (x_{\alpha_n} - x^\dagger) \\ &= \sum_{m\in\mathbb{N}} \frac{\sigma_m}{\sigma_m^2 + \alpha_n}(\delta_n u_n, u_m)_Y\, v_m + (x_{\alpha_n} - x^\dagger) \\ &= \frac{\delta_n\sigma_n}{\sigma_n^2+\alpha_n}v_n + (x_{\alpha_n} - x^\dagger). \end{aligned}$$

Aus Annahme (6.3) für $y^\delta = y_n$ sowie für $y^\delta = y$ erhalten wir daraus

$$\frac{\sigma_n\delta_n^{1/3}}{\sigma_n^2+\alpha_n} \leq \|x^{\delta_n}_{\alpha_n} - x^\dagger\|_X\delta_n^{-2/3} + \|x_{\alpha_n} - x^\dagger\|_X\delta_n^{-2/3} \to 0 \qquad \text{für } n \to \infty.$$

Andererseits folgt aber aus $\sigma_n = \delta_n^{1/3}$ und $\alpha_n\delta_n^{-2/3} \to 0$ auch

$$\frac{\sigma_n\delta_n^{1/3}}{\sigma_n^2+\alpha_n} = \frac{\delta_n^{2/3}}{\delta_n^{2/3}+\alpha_n} = \frac{1}{1+\alpha_n\delta_n^{-2/3}} \to 1 \qquad \text{für } n \to \infty$$

und damit der gesuchte Widerspruch. □

Ein Vergleich der Darstellung (6.1) mit den Normalengleichungen (3.6) legt nahe, dass die Tikhonov-Regularisierung auch eine Minimierungseigenschaft hat. Dies trifft in der Tat zu.

Satz 6.5. *Für $y \in Y$ ist $x_\alpha = R_\alpha y$ eindeutig bestimmt als Minimierer des* Tikhonov-Funktionals

$$J_\alpha(x) := \frac{1}{2}\|Kx - y\|_Y^2 + \frac{\alpha}{2}\|x\|_X^2. \tag{6.4}$$

Beweis. Ein Minimierer $\bar{x} \in X$ von J_α ist charakterisiert durch $J_\alpha(\bar{x}) \le J_\alpha(x)$ für alle $x \in X$. Wir betrachten daher für beliebiges $x \in X$ und eine Lösung x_α von (6.1) die Differenz der Funktionalwerte, sortieren die Skalarprodukte um, und erhalten

$$\begin{aligned} J_\alpha(x) - J_\alpha(x_\alpha) &= \frac{1}{2}(Kx - y, Kx - y)_Y + \frac{\alpha}{2}(x, x)_X - \frac{1}{2}(Kx_\alpha - y, Kx_\alpha - y)_Y - \frac{\alpha}{2}(x_\alpha, x_\alpha)_X \\ &= \frac{1}{2}\|Kx - Kx_\alpha\|_Y^2 + \frac{\alpha}{2}\|x - x_\alpha\|_X^2 + (K^*(Kx_\alpha - y) + \alpha x_\alpha, x - x_\alpha)_X \\ &= \frac{1}{2}\|Kx - Kx_\alpha\|_Y^2 + \frac{\alpha}{2}\|x - x_\alpha\|_X^2 \\ &\ge 0, \end{aligned}$$

wobei wir im vorletzten Schritt (6.1) verwendet haben. Also ist x_α ein Minimierer von J_α.

Gilt umgekehrt $J_\alpha(x) - J_\alpha(\bar{x}) \ge 0$ für alle $x \in X$, so folgt insbesondere für $x = \bar{x} + tz$ für $t > 0$ beliebig und $z \in X$ fest

$$0 \le J_\alpha(\bar{x} + tz) - J_\alpha(\bar{x}) = \frac{t^2}{2}\|Kz\|_Y^2 + \frac{t^2\alpha}{2}\|z\|_X^2 + t\,(K^*(K\bar{x} - y) + \alpha\bar{x}, z)_X\,.$$

Division durch t und Grenzübergang $t \to 0$ ergibt also

$$(K^*(K\bar{x} - y) + \alpha\bar{x}, z)_X \ge 0 \qquad \text{für alle } z \in X.$$

Dies ist aber nur dann möglich, wenn $K^*K\bar{x} + \alpha\bar{x} = K^*y$ gilt. Da x_α die eindeutige Lösung von (6.1) ist, folgt $\bar{x} = x_\alpha$. Also ist x_α der eindeutige Minimierer von (6.4). □

Die Charakterisierung der Tikhonov-Regularisierung als Minimierung des Funktionals (6.4) stellt einen weiteren Zusammenhang zur Minimum-Norm-Lösung $x^\dagger$ her: Anstelle auf einer reinen Ausgleichslösung zu bestehen, deren Norm für $y \notin \mathcal{D}(K^\dagger)$ nicht beschränkt sein muss, wird eine Näherungslösung gesucht, die gleichzeitig die (quadrierte) *Residuumsnorm* $\|Kx - y\|_Y^2$ und die (quadrierte) Norm $\|x\|_X^2$ minimiert.[1] Der Regularisierungsparameter α bestimmt dabei die Gewichtung: je kleiner das Fehlerniveau δ ist, desto mehr Gewicht kann man der Minimierung der Residuumsnorm geben (d. h. desto kleiner kann man α wählen).

[1] In dieser Form wurde diese Regularisierung auch von Andreĭ Nikolaevich Tikhonov, einem bedeutenden russischen Mathematiker des 20. Jahrhunderts, eingeführt; siehe [Tikhonov 1963a; Tikhonov 1963b].

Umgekehrt verlangt ein höheres Fehlerniveau eine stärkere Gewichtung des *Strafterms* $\|x\|_X^2$ (und damit ein größeres α), damit die Näherung stabil bleibt.

Darüberhinaus liefert Satz 6.5 eine neue Interpretation der einfachsten Quellbedingung $x^\dagger \in X_1 = \mathcal{R}(K^*)$. Der Minimierer von (6.4) ändert sich nicht, wenn wir das Funktional durch $\alpha > 0$ dividieren; also ist x_α^δ auch gegeben als die Lösung von

$$\min_{x\in X} \frac{1}{2\alpha}\|Kx - y^\delta\|_Y^2 + \frac{1}{2}\|x\|_X^2. \tag{6.5}$$

Für $\delta \to 0$ und $\alpha \to 0$ soll nun $x_\alpha^\delta \to x^\dagger$ konvergieren. Machen wir (formal) den Grenzübergang in (6.5), d. h. ersetzen wir zuerst y^δ durch $y \in \mathcal{R}(K)$ und lassen $\alpha \to 0$ gehen, so kann das Funktional nur dann einen endliches Minimum in $\bar{x}$ annehmen, falls $K\bar{x} = y$ gilt. Im Grenzfall $\alpha = 0$ erhalten wir also das Funktional

$$\min_{x\in X,\ Kx=y} \frac{1}{2}\|x\|_X^2. \tag{6.6}$$

Wir gehen weiter formal vor: Mit Hilfe eines Lagrange-Multiplikators $p \in Y$ kann (6.6) als unbeschränktes Sattelpunktproblem

$$\min_{x\in X}\max_{p\in Y} \frac{1}{2}\|x\|_X^2 - (p, Kx - y)_Y$$

geschrieben werden. Damit $(\bar{x}, \bar{p}) \in X \times Y$ ein Sattelpunkt sein kann, müssen dort die Ableitungen nach x und p verschwinden: wir erhalten die beiden Bedingungen.

$$\begin{cases} \bar{x} = K^*\bar{p}, \\ K\bar{x} = y. \end{cases}$$

Für $y \in \mathcal{R}(K)$ beschreibt die Lösung von (6.6) aber genau die Minimum-Norm-Lösung $x^\dagger$, d. h. $\bar{x} = x^\dagger$. Die Existenz eines Lagrange-Multiplikators $\bar{p}$ mit $x^\dagger = K^*\bar{p}$ entspricht daher genau der Quellbedingung $x^\dagger \in \mathcal{R}(K^*)$. (Da K^* nicht surjektiv sein muss, ist dies eine nichttriviale Forderung.) Anschaulich ist dies nachvollziehbar: Wenn wir $x^\dagger$ über eine Folge von Minimierern x_α^δ annähern möchten, so sollte $x^\dagger$ selbst ein Minimierer (eines geeigneten Grenzproblems) sein.

Diese Interpretation der Tikhonov-Regularisierung als Minimierung eines Funktionals lässt sich – im Gegensatz zur Spektraldarstellung – auf nichtlineare Operatorgleichung erweitern. Sie ist auch in allgemeinen Banachräumen anwendbar und lässt sich sogar durch Verwendung anderer Diskrepanz- und Strafterme als Normen weiter verallgemeinern. (Dies gilt auch für die entsprechende Interpretation der Quellbedingung.) Dann sind natürlich andere Beweismethoden und insbesondere andere Quellbedingungen notwendig.

7 LANDWEBER-REGULARISIERUNG

Ausgangspunkt für die Landweber-Regularisierung ist die Charakterisierung der Minimum-Norm-Lösung $x^\dagger$ nach Folgerung 3.6 als Lösung $x \in \mathcal{N}(K)^\perp$ der Normalengleichung (3.6). Diese können äquivalent geschrieben werden für beliebiges $\omega > 0$ als Fixpunktgleichung

$$x = x - \omega(K^*Kx - K^*y) = x + \omega K^*(y - Kx).$$

Die zugehörige Fixpunktiteration – auch als *Richardson-Iteration*[1] bekannt – lautet

$$x_n = x_{n-1} + \omega K^*(y - Kx_{n-1}), \qquad n \in \mathbb{N}, \tag{7.1}$$

für ein $x_0 \in X$. Wir werden hier nur $x_0 = 0$ betrachten. Aus dem Banachschen Fixpunktsatz folgt, dass diese Iteration gegen eine Lösung der Normalengleichung konvergiert, falls $y \in \mathcal{R}(K)$ und $\|\operatorname{Id} - \omega K^*K\|_{\mathcal{L}(X,X)} < 1$ ist. Wegen $x_0 = 0 \in \mathcal{R}(K^*)$ ist auch $x_n \in \mathcal{R}(K^*) \subset \mathcal{N}(K)^\perp$, und damit konvergiert x_n gegen $x^\dagger$. Für $y^\delta \notin \mathcal{R}(K)$ kann man hingegen keine Konvergenz erwarten. Die Idee ist nun, die Iteration rechtzeitig abzubrechen, d. h. x_m für ein geeignetes $m \in \mathbb{N}$ als regularisierte Näherung zu akzeptieren. Hier spielt also der *Abbruchindex* $m \in \mathbb{N}$ die Rolle des Regularisierungsparameters, was mit $\alpha = \frac{1}{m} > 0$ zu der Schreibweise der letzten Kapitel passt.[2]

Die nach m Schritten abgebrochene Iteration (7.1) kann in die Form einer spektralen Regularisierung gebracht werden. Dafür leiten wir zuerst eine rekursionsfreie Darstellung letzten Iterierten x_m her.

Lemma 7.1. *Für $x_0 = 0$ gilt*

$$x_m = \omega \sum_{n=0}^{m-1} (\operatorname{Id} - \omega K^*K)^n K^*y \qquad \textit{für alle } m \in \mathbb{N}.$$

[1] Diese Methode zur Lösung von linearen Gleichungssystemen geht auf Lewis Fry Richardson zurück. Er propagierte auch 1922 die heutige Methode der Wettervorhersage auf Basis von numerischer Simulation. (Ein eigener erster Versuch von 1910 – durchgeführt von Hand! – war grundsätzlich korrekt, lieferte aber wegen gestörter Eingabedaten ein falsches Ergebnis. Wettervorhersage ist ein schlecht gestelltes Problem!)

[2] Zur Lösung von schlecht gestellten Operatorgleichungen wurde diese Methode zuerst von Lawrence Landweber betrachtet. In [Landweber 1951] zeigt er die Konvergenz für $y \in \mathcal{R}(K)$; andernfalls, schreibt er dort, könnten die Iterierten „als nützliche Näherungslösung dienen“.

Beweis. Dies folgt mit vollständiger Induktion: Für $m = 1$ ist

$$x_1 = \omega K^* y = \omega(\mathrm{Id} - \omega K^* K)^0 K^* y.$$

Sei nun $m \in \mathbb{N}$ beliebig und gelte die behauptete Darstellung für x_m. Dann ist

$$\begin{aligned} x_{m+1} &= x_m + \omega K^*(y - Kx_m) \\ &= (\mathrm{Id} - \omega K^* K)x_m + \omega K^* y \\ &= (\mathrm{Id} - \omega K^* K)\left(\omega \sum_{n=0}^{m-1} (\mathrm{Id} - \omega K^* K)^n K^* y\right) + \omega K^* y \\ &= \omega \sum_{n=0}^{m-1} (\mathrm{Id} - \omega K^* K)^{n+1} K^* y + \omega(\mathrm{Id} - \omega K^* K)^0 K^* y \\ &= \omega \sum_{n=0}^{m} (\mathrm{Id} - \omega K^* K)^n K^* y. \end{aligned}$$

□

Die nach m Iterationen abgebrochene *Landweber-Iteration* (7.1) wird also erzeugt durch einen linearen Operator, d. h.

$$x_m = \varphi_m(K^* K) K^* y$$

mit

$$\varphi_m(\lambda) = \omega \sum_{n=0}^{m-1} (1 - \omega\lambda)^n = \omega \frac{1 - (1 - \omega\lambda)^m}{1 - (1 - \omega\lambda)} = \frac{1 - (1 - \omega\lambda)^m}{\lambda}.$$

Bis auf die Schreibweise φ_m anstelle von φ_α für $\alpha = \frac{1}{m}$ (d. h. statt $\alpha \to 0$ betrachten wir $m \to \infty$) hat dieser Filter genau die Form aus Beispiel 5.2 (iii).

Satz 7.2. *Für $\omega \in (0, \kappa^{-1})$ wird durch $\{\varphi_m\}_{m\in\mathbb{N}}$ eine Regularisierung $\{R_m\}_{m\in\mathbb{N}}$ mit $R_m := \varphi_m(K^*K)K^*$ definiert.*

Beweis. Wir müssen lediglich zeigen, dass $\varphi_m(\lambda) \to \frac{1}{\lambda}$ für $m \to \infty$ konvergiert und $\lambda\varphi_m(\lambda)$ gleichmäßig beschränkt ist für alle $m \in \mathbb{N}$. Wegen der Bedingung an ω gilt $0 < 1 - \omega\lambda < 1$ für alle $\lambda \in (0, \kappa]$, woraus sowohl $(1 - \omega\lambda)^m \to 0$ für $m \to \infty$ als auch

$$\lambda|\varphi_m(\lambda)| = |1 - (1 - \omega\lambda)^m| \le 1 =: C_\varphi \qquad \text{für alle } m \in \mathbb{N} \text{ und } \lambda \in (0, \kappa]$$

folgt. Also ist $\{\varphi_m\}_{m\in\mathbb{N}}$ ein regularisierender Filter, und Satz 5.6 liefert die Behauptung. □

Die Landweber-Iteration konvergiert also genau dann für $m \to \infty$ gegen die Minimum-Norm-Lösung $x^\dagger$, wenn $y \in \mathcal{D}(K^\dagger)$ liegt; ansonsten divergiert sie. Es liegt nun nahe, den Abbruchindex nach dem Diskrepanzprinzip zu wählen: Zu $\tau > 1$ bestimme $m(\delta, y^\delta)$ so, dass für $x_m^\delta := R_m y^\delta$ gilt

$$\|Kx^\delta_{m(\delta,y^\delta)} - y^\delta\|_Y \le \tau\delta < \|Kx_m^\delta - y^\delta\|_Y \qquad \text{für alle } m < m(\delta, y^\delta). \tag{7.2}$$

(Dies ist kein zusätzlicher Rechenaufwand, denn das Residuum $y^\delta - Kx_m^\delta$ wird als Teil der Iterationsvorschrift (7.1) berechnet.) Die Existenz solch eines $m(\delta, y^\delta)$ garantiert dabei Satz 4.7 zusammen mit Lemma 5.3. Im Folgenden sei stets $\omega \in (0, \kappa^{-1})$ angenommen.

Satz 7.3. *Für alle $\nu > 0$ und $\tau > 1$ ist die Landweber-Regularisierung zusammen mit der Parameterwahlstrategie (7.2) ein ordnungsoptimales Regularisierungsverfahren.*

Beweis. Wir weisen die notwendigen Eigenschaften für φ_m nach, wobei wir $\alpha := \frac{1}{m}$ setzen. Zuerst gilt wegen $\omega\lambda < 1$ mit der Bernoullischen Ungleichung

$$|\varphi_m(\lambda)| = \frac{|1-(1-\omega\lambda)^m|}{\lambda} \le \frac{|1-1+m\omega\lambda|}{\lambda} = \omega m = \omega\alpha^{-1} \qquad \text{für alle } \lambda \in (0,\kappa]$$

und damit die Bedingung (5.7) (falls $\omega \le 1$; ansonsten zeigt die Betrachtung des Beweises von Satz 5.11, dass dadurch nur die Konstante C_2 vergrössert wird).

Weiter folgt aus der Bernoullischen Ungleichung auch $(1+x) \le e^x$ und damit

$$r_m(\lambda) = 1 - \lambda\varphi_m(\lambda) = (1-\omega\lambda)^m \le e^{-\omega\lambda m} \le 1 =: C_r \qquad \text{für alle } m \in \mathbb{N}, \lambda \in (0,\kappa].$$

Wir betrachten nun für $\nu > 0$ und m fest die Funktion $h_m(\lambda) := \lambda^{\nu/2} e^{-\omega\lambda m}$ und berechnen

$$h_m'(\lambda) = \frac{\nu}{2}\lambda^{\nu/2-1}e^{-\omega\lambda m} - \omega m\lambda^{\nu/2}e^{-\omega\lambda m} = \lambda^{\nu/2-1}e^{-\omega\lambda m}\omega m\left(\frac{\nu}{2\omega m} - \lambda\right).$$

In der Nullstelle $\lambda^* = \frac{\nu}{2\omega m}$ ist $h_m''(\lambda^*) < 0$ und daher

$$\sup_{\lambda\in(0,\kappa]} \lambda^{\nu/2} r_m(\lambda) \le \sup_{\lambda\in(0,\infty)} h_m(\lambda) = h_m\left(\frac{\nu}{2\omega m}\right) = e^{-\nu/2}\left(\frac{\nu}{2\omega}\right)^{\nu/2} m^{-\nu/2} =: C_\nu \alpha^{\nu/2}.$$

Damit gilt (5.8) für alle $\nu > 0$. Die Landweber-Iteration hat also unendliche Qualifikation, und die Aussage folgt für $\tau > C_r = 1$ aus Satz 5.11. □

Wir untersuchen nun die Monotonie-Eigenschaften der Landweber-Iteration.

Satz 7.4. *Sei $m \in \mathbb{N}$. Gilt $Kx_m^\delta - y^\delta \neq 0$, so ist*

$$\|Kx_{m+1}^\delta - y^\delta\|_Y < \|Kx_m^\delta - y^\delta\|_Y.$$

Beweis. Aus der Iterationsvorschrift (7.1) folgt

$$\begin{aligned} Kx_{m+1}^\delta - y^\delta &= K\left((\mathrm{Id} - \omega K^*K)x_m^\delta + \omega K^* y^\delta\right) - y^\delta \\ &= (\mathrm{Id} - \omega KK^*)Kx_m^\delta - (\mathrm{Id} + \omega KK^*)y^\delta \\ &= (\mathrm{Id} - \omega KK^*)(Kx_m^\delta - y^\delta) \end{aligned}$$

und damit wegen $\omega < \kappa^{-1} = \sigma_1^{-2} \leq \sigma_n^{-2}$ für alle $n \in \mathbb{N}$

$$\begin{aligned}\|Kx_{m+1}^\delta - y^\delta\|_Y^2 &= \sum_{n\in\mathbb{N}} (1-\omega\sigma_n^2)^2 \left|\left(Kx_m^\delta - y^\delta, u_n\right)_Y\right|^2 \\ &< \sum_{n\in\mathbb{N}} \left|\left(Kx_m^\delta - y^\delta, u_n\right)_Y\right|^2 \leq \|Kx_m^\delta - y^\delta\|_Y^2. \end{aligned}$$ □

Das Residuum ist also stets monoton fallend. Für den Fehler gilt dies aber nur bedingt.

Satz 7.5. *Sei $m \in \mathbb{N}$. Gilt*

$$\|Kx_m^\delta - y^\delta\|_Y > 2\delta,$$

so ist

$$\|x_{m+1}^\delta - x^\dagger\|_X < \|x_m^\delta - x^\dagger\|_X.$$

Beweis. Wir verwenden wieder die Iterationsvorschrift und schreiben mit $\rho_m^\delta = y^\delta - Kx_m^\delta$ und $y = Kx^\dagger$

$$\begin{aligned}\|x_{m+1}^\delta - x^\dagger\|_X^2 &= \|x_m^\delta - x^\dagger + \omega K^*(y^\delta - Kx_m^\delta)\|_X^2 \\ &= \|x_m^\delta - x^\dagger\|_X^2 - 2\omega\left(Kx^\dagger - Kx_m^\delta, \rho_m^\delta\right)_Y + \omega^2\|K^*\rho_m^\delta\|_X^2 \\ &= \|x_m^\delta - x^\dagger\|_X^2 + \omega\left(\rho_m^\delta - 2y + 2Kx_m^\delta, \rho_m^\delta\right)_Y + \omega\left(\omega\|K^*\rho_m^\delta\|_X^2 - \|\rho_m^\delta\|_Y^2\right).\end{aligned}$$

Wir müssen nun zeigen, dass die letzten beiden Terme negativ sind. Für den ersten Term verwenden wir die Definition von ρ_m^δ und erhalten durch Einsetzen von $\rho_m^\delta = 2\rho_m^\delta - \rho_m^\delta = 2y^\delta - 2Kx_m^\delta - \rho_m^\delta$, dass gilt

$$\begin{aligned}\left(\rho_m^\delta - 2y + 2Kx_m^\delta, \rho_m^\delta\right)_Y &= 2\left(y^\delta - y, \rho_m^\delta\right)_Y - \|\rho_m^\delta\|_Y^2 \\ &\leq 2\delta\|\rho_m^\delta\|_Y - \|\rho_m^\delta\|_Y^2 \\ &= \left(2\delta - \|Kx_m^\delta - y^\delta\|_Y\right)\|\rho_m^\delta\|_Y < 0,\end{aligned}$$

da nach Annahme $\|\rho_m^\delta\|_Y > 2\delta > 0$ und damit die Klammer negativ ist. Für den zweiten Term in Klammern verwenden wir, dass wegen $\omega < \kappa^{-1}$ gilt

$$\omega\|K^*\rho_m^\delta\|_X^2 \leq \omega\|K^*\|_{\mathcal{L}(Y,X)}^2\|\rho_m^\delta\|_Y^2 = \omega\kappa\|\rho_m^\delta\|^2 < \|\rho_m^\delta\|_X^2$$

und damit

$$\omega\left(\omega\|K^*\rho_m^\delta\|_X^2 - \|\rho_m^\delta\|_Y^2\right) < 0. \tag{7.3}$$

Zusammen erhalten wir die gewünschte Ungleichung. □

Die Landweber-Iteration reduziert also zunächst den Fehler, bis das Residuum unter das doppelte Fehlerniveau fällt. (Für das Diskrepanzprinzip sollte daher stets $\tau \leq 2$ gewählt werden, um nicht garantiert zu früh abzubrechen.) Danach wird für $y^\delta \notin \mathcal{R}(K)$ aufgrund von Satz 5.6 der Fehler aber wieder anwachsen. Dieses Verhalten wird *Semikonvergenz* genannt und ist typisch für iterative Verfahren, wenn sie auf schlecht gestellte Probleme angewendet werden. Das Diskrepanzprinzip verhindert dann, dass der Fehler wieder beliebig anwächst. (Ein leichtes Anwachsen wird dabei in Kauf genommen – wie stark, hängt von der Wahl von $\tau \in (1, 2)$ ab.) Der folgende Satz gibt eine obere Schranke für die Anzahl der dafür notwendigen Schritte an.

Satz 7.6. *Für $\tau > 1$ und $y^\delta \in B_\delta(Kx^\dagger)$ bricht das Diskrepanzprinzip ab im Schritt*

$$m(\delta, y^\delta) \leq C\delta^{-2} \qquad \text{für ein } C > 0.$$

Beweis. Wir leiten zuerst eine Konvergenzrate für das Residuum in Abhängigkeit von m her. Dazu betrachten wir für $n \geq 0$ die Iterierten x_n, die die Landweber-Iteration mit den exakten Daten $y := Kx^\dagger \in \mathcal{R}(K)$ erzeugt, und bezeichnen das entsprechende Residuum mit $\rho_n := y - Kx_n$. Wir schätzen nun ähnlich wie im Beweis von Satz 7.5 ab. Aus der Iterationsvorschrift folgt durch einfaches Umformen und mit (7.3)

$$\begin{aligned}
\|x^\dagger - x_n\|_X^2 - \|x^\dagger - x_{n+1}\|_X^2 &= \|x^\dagger - x_n\|_X^2 - \|x^\dagger - x_n - \omega K^* \rho_n\|_X^2 \\
&= 2\omega \left(Kx^\dagger - Kx_n, \rho_n\right)_Y - \omega^2 \|K^* \rho_n\|_X^2 \\
&= \omega \left(\|\rho_n\|_Y^2 - \omega \|K^* \rho_n\|_X^2\right) + \omega \|\rho_n\|_Y^2 \\
&> \omega \|\rho_n\|_Y^2.
\end{aligned}$$

Summieren über $n = 0, \ldots, m-1$ ergibt zusammen mit der Monotonie des Residuums aus Satz 7.4, dass gilt

$$\begin{aligned}
\|x^\dagger - x_0\|_X^2 - \|x^\dagger - x_m\|_X^2 &= \sum_{n=0}^{m-1} \left(\|x^\dagger - x_n\|_X^2 - \|x^\dagger - x_{n+1}\|_X^2\right) \\
&> \omega \sum_{n=0}^{m-1} \|\rho_n\|_Y^2 > \omega m \|\rho_m\|_X^2.
\end{aligned}$$

Insbesondere ist daher

$$\|y - Kx_m\|_Y^2 < (\omega m)^{-1} \|x^\dagger - x_0\|_X^2.$$

Wie im Beweis von Satz 7.4 gilt nun wegen $x_0 = 0$

$$\rho_m^\delta = y^\delta - Kx_m^\delta = (\mathrm{Id} - \omega KK^*)(y^\delta - Kx_{m-1}^\delta) = \cdots = (\mathrm{Id} - \omega KK^*)^m y^\delta$$

und analog $\rho_m = (\mathrm{Id} - \omega KK^*)^m y$. Daraus folgt wegen $\omega < \kappa^{-1} < \sigma_n^{-2}$ die Abschätzung

$$\|(\mathrm{Id} - \omega KK^*)^m (y^\delta - y)\|_Y^2 = \sum_{n\in\mathbb{N}} (1 - \omega\sigma_n^2)^{2m} \left| \left(y^\delta - y, u_n \right)_Y \right|^2 \leq \|y^\delta - y\|_Y^2$$

und damit

$$\begin{aligned} \|Kx_m^\delta - y^\delta\|_Y &= \|(\mathrm{Id} - \omega KK^*)^m y^\delta\|_Y \\ &\leq \|(\mathrm{Id} - \omega KK^*)^m y\|_Y + \|(\mathrm{Id} - \omega KK^*)^m (y^\delta - y)\|_Y \\ &\leq \|y - Kx_m\|_Y + \|y^\delta - y\|_Y \\ &\leq (\omega m)^{-1/2} \|x^\dagger - x_0\|_X + \delta. \end{aligned}$$

Das Diskrepanzprinzip bestimmt nun den Abbruchindex $m(\delta, y^\delta)$ so, dass zum ersten Mal $\|Kx_{m(\delta,y^\delta)}^\delta - y^\delta\|_Y \leq \tau\delta$ ist. Wegen der Monotonie des Residuums gilt das spätestens für das erste $\bar{m} \in \mathbb{N}$ mit

$$(\omega\bar{m})^{-1/2} \|x^\dagger - x_0\|_X + \delta \leq \tau\delta$$

beziehungsweise

$$\bar{m} \geq \omega \frac{\|x^\dagger - x_0\|_X^2}{\omega^2(\tau - 1)^2} \delta^{-2} \geq \bar{m} - 1.$$

Es ist also sicher

$$m(\delta, y^\delta) \leq \bar{m} - 1 \leq C\delta^{-2} + 1$$

mit $C := \omega^{-1}(\tau - 1)^{-2} \|x^\dagger - x_0\|_X^2 + \max\{1, \delta^2\}$. □

Es überrascht nicht, dass man diese Abschätzung unter der üblichen Quellbedingung $x^\dagger \in X_\nu$ noch verbessern kann.

Folgerung 7.7. *Ist $x^\dagger \in X_\nu$ für ein $\nu > 0$, so bricht das Diskrepanzprinzip für $\tau > 1$ und $y^\delta \in B_\delta(Kx^\dagger)$ ab im Schritt*

$$m(\delta, y^\delta) \leq C\delta^{-\frac{2}{\nu+1}} \qquad \text{für ein } C > 0.$$

Beweis. Aus (5.16) im Beweis von Satz 5.11 folgt mit $\alpha = \frac{1}{m}$ die Abschätzung

$$m(\delta, y^\delta)^{1/2} = \alpha(\delta, y^\delta)^{-1/2} \leq C\delta^{-\frac{1}{\nu+1}}.$$ □

Trotzdem erfordert die Landweber-Regularisierung in der Praxis oft zu viele Iterationen, bis das Abbruchkriterium erreicht ist, und man greift zu beschleunigten Varianten wie etwa in [Engl, Hanke & Neubauer 1996, Kapitel 6.2, 6.3] beschrieben. Außerdem lässt sich der Ansatz, ein iteratives Verfahren zur Lösung der Normalengleichung durch vorzeitigen Abbruch in ein Regularisierungsverfahren zu verwandeln, auch auf andere Verfahren als die Richardson-Iteration anwenden; eine beliebte Wahl ist dabei das Verfahren der konjugierten Gradienten; siehe [Engl, Hanke & Neubauer 1996, Kapitel 7].

8 DISKRETISIERUNG ALS REGULARISIERUNG

Und nun zu etwas völlig anderem. Wie wir gesehen haben, liegt die fundamentale Schwierigkeit bei inversen Problemen in der Unstetigkeit der Pseudoinversen für kompakte Operatoren $K : X \to Y$ mit unendlichdimensionalem Bild. Es ist daher naheliegend zu versuchen, eine Folge von Operatoren K_n mit *endlichdimensionalem* Bild zu konstruieren und die gesuchte Minimum-Norm-Lösung $K^\dagger y$ mit Hilfe der (nun stetigen) Pseudoinversen $(K_n)^\dagger$ anzunähern. Dies funktioniert – bis zu einem gewissen Grad – tatsächlich. Um ein endlichdimensionales Bild zu erhalten, haben wir im wesentlichen zwei Möglichkeiten:

(i) Wir schränken das Urbild von K auf einen endlichdimensionalen Unterraum $X_n \subset X$ ein und definieren $K_n : X_n \to Y$. (Ist nämlich $\{x_1, \dots, x_n\}$ eine Basis von X_n, so ist auch $\mathcal{R}(K_n) = \operatorname{span}\{Kx_1, \dots, Kx_n\}$ endlichdimensional.) Man bezeichnet diesen Ansatz als *Ausgleichsprojektion* (englisch: „least-squares projection“).

(ii) Wir schränken direkt das Bild von K auf einen endlichdimensionalen Unterraum $Y_n \subset Y$ ein und definieren $K_n : X \to Y_n$. Dies bezeichnet man als *duale Ausgleichsprojektion* (englisch: „dual least-squares projection“).

Natürlich kann man auch Bild *und* Urbild einschränken und $K_n : X_n \to Y_n$ definieren; dies bietet aber aus Sicht der Regularisierung keinen Vorteil. Wir betrachten nun beide Ansätze, wobei sich der zweite als vorteilhafter herausstellen wird. Da wir dabei keine Spektraltheorie benötigen, betrachten wir hier wieder einen lediglich beschränkten Operator $T \in \mathcal{L}(X, Y)$.

8.1 AUSGLEICHSPROJEKTION

Sei eine Folge $\{X_n\}_{n\in\mathbb{N}}$ von geschachtelten Unterräumen

$$X_1 \subset X_2 \subset \cdots \subset X$$

mit $\dim X_n = n$ und $\overline{\bigcup_{n\in\mathbb{N}} X_n} = X$ gegeben. Sei weiterhin $P_n := P_{X_n}$ die orthogonale Projektion auf X_n und $T_n := TP_n \in \mathcal{L}(X, Y)$. Da T_n endlichdimensionales Bild hat, ist $T_n^\dagger := (T_n)^\dagger$ stetig. Wir wählen also für $y \in Y$ die Regularisierung $x_n := T_n^\dagger y$, d. h. die Minimum-Norm-Lösung von $TP_n x = y$. Nach Lemma 3.4 ist

$$x_n \in \mathcal{R}(T_n^\dagger) = \mathcal{N}(T_n)^\perp = \overline{\mathcal{R}(T_n^*)} = \overline{\mathcal{R}(P_n T^*)} \subset X_n,$$

da X_n endlichdimensional und damit abgeschlossen und P_n selbstadjungiert ist. (Wir suchen also nach der Minimum-Norm-Lösung nur in X_n statt in ganz X.) Um zu zeigen, dass $T_n^\dagger$ eine Regularisierung im Sinne von Definition 4.1 darstellt, müssen wir nun zeigen, dass für $y \in \mathcal{D}(T^\dagger)$ gilt $x_n \to x^\dagger$. Dies können wir über die Norm von x_n charakterisieren.[1]

Lemma 8.1. *Für $y \in \mathcal{D}(T^\dagger)$ konvergiert $x_n \to x^\dagger$ genau dann, wenn gilt* $\limsup_{n\to\infty} \|x_n\|_X \leq \|x^\dagger\|_X$.

Beweis. Aus $x_n \to x^\dagger$ folgt sofort mit der Dreiecksungleichung

$$\|x_n\|_X \leq \|x_n - x^\dagger\|_X + \|x^\dagger\|_X \to \|x^\dagger\|_X.$$

Ist umgekehrt die lim sup-Bedingung erfüllt, so ist die Folge $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ und damit auch $\{x_n\}_{n\in\mathbb{N}}$ in X beschränkt. Es existiert also eine schwach konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ und ein $\bar{x} \in X$ mit $x_k := x_{n_k} \rightharpoonup \bar{x}$ und $Tx_k \rightharpoonup T\bar{x}$. Nach Definition von x_k als Ausgleichslösung von $T_k x = y$ (mit minimaler Norm) sowie $Tx^\dagger = P_{\overline{\mathcal{R}}} y$ nach Lemma 3.4 (iv) und $\overline{\mathcal{R}(T)} = \mathcal{N}(T)^\perp$ gilt nun

$$\begin{aligned}\|T_k x_k - Tx^\dagger\|_Y^2 + \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})y\|_Y^2 &= \|T_k x_k - y\|_Y^2 \leq \|T_k x - y\|_Y^2 \\ &= \|T_k x - Tx^\dagger\|_Y^2 + \|(\mathrm{Id} - P_{\overline{\mathcal{R}}})y\|_Y^2 \qquad \text{für alle } x \in X.\end{aligned}$$

Wegen $x_k \in X_k$, d. h. $x_k = P_k x_k$, folgt daraus für $x = P_k x^\dagger$

$$\begin{aligned}\|Tx_k - Tx^\dagger\|_Y = \|T_k x_k - Tx^\dagger\|_Y &\leq \|T_k P_k x^\dagger - Tx^\dagger\|_Y = \|TP_k x^\dagger - Tx^\dagger\|_Y \\ &\leq \|T\|_{\mathcal{L}(X,Y)} \|(\mathrm{Id} - P_k)x^\dagger\|_X.\end{aligned} \tag{8.1}$$

Nach Annahme an $\{X_n\}_{n\in\mathbb{N}}$ konvergiert der letzte Term gegen Null für $k \to \infty$, und daher gilt $Tx_k \to Tx^\dagger$. Also muss $\bar{x} - x^\dagger \in \mathcal{N}(T)$ sein. Nun ist stets $x^\dagger \in \mathcal{N}(T)^\perp$, und daher folgt aus der schwachen Unterhalbstetigkeit der Norm zusammen mit der lim sup-Bedingung

$$\|\bar{x} - x^\dagger\|_X^2 + \|x^\dagger\|_X^2 = \|\bar{x}\|_X^2 \leq \liminf_{k\to\infty} \|x_k\|_X^2 \leq \limsup_{k\to\infty} \|x_k\|_X^2 \leq \|x^\dagger\|_X^2$$

und damit $\bar{x} = x^\dagger$. Jede Teilfolge konvergiert daher schwach gegen den selben Grenzwert $x^\dagger$. Also muss die gesamte Folge (schwach) gegen $x^\dagger$ konvergieren.

Schließlich ergibt die Unterhalbstetigkeit zusammen mit der lim sup-Bedingung, dass auch $\|x_n\|_X \to \|x^\dagger\|_X$ konvergiert, woraus zusammen mit der schwachen Konvergenz die starke Konvergenz folgt. □

[1] Wir folgen hier [Kindermann 2016]; der entsprechende Beweis in [Engl, Hanke & Neubauer 1996] über eine analoge Aussage für die schwache Konvergenz ist fehlerhaft und benötigt eine zusätzliche Annahme, wie in [Du 2008] bemerkt wurde.

Leider lassen sich Beispiele konstruieren, in denen $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ nicht beschränkt ist; siehe zum Beispiel [Engl, Hanke & Neubauer 1996, Example 3.19]. Eine hinreichende Bedingung für die Konvergenz gibt der folgende Satz.

Satz 8.2. *Sei $y \in \mathcal{D}(T^\dagger)$ und*

$$\limsup_{n\to\infty} \|(T_n^*)^\dagger x_n\|_Y = \limsup_{n\to\infty} \|(T_n^\dagger)^* x_n\|_Y < \infty. \tag{8.2}$$

Dann konvergiert $x_n \to x^\dagger$.

Beweis. Wegen

$$\|x_n\|_X^2 = \left(x_n - x^\dagger, x_n\right)_X + \left(x^\dagger, x_n\right)_X \le \left(x_n - x^\dagger, x_n\right)_X + \|x^\dagger\|_X \|x_n\|_X$$

genügt es zu zeigen, dass der erste Term auf der rechten Seite gegen Null konvergiert. Dafür setzen wir $w_n := (T_n^\dagger)^* x_n$ und verwenden, dass wegen $x_n \in \mathcal{R}(T_n^\dagger) = \mathcal{R}(T_n^*)$ (denn $\mathcal{R}(T_n^*) \subset X_n$ ist endlichdimensional) gilt $T_n^* w_n = x_n$. Wir können daher wie folgt abschätzen:

$$\begin{aligned}
\left(x_n - x^\dagger, x_n\right)_X &= \left(x_n - x^\dagger, T_n^* w_n\right)_X = \left(T_n x_n - T_n x^\dagger, w_n\right)_Y \\
&= \left(T_n x_n - T x^\dagger, w_n\right)_Y + \left(T x^\dagger - T_n x^\dagger, w_n\right)_Y \\
&\le \left(\|T_n x_n - T x^\dagger\|_Y + \|T(\mathrm{Id} - P_n) x^\dagger\|_Y\right) \|w_n\|_Y \\
&\le 2\|T\|_{\mathcal{L}(X,Y)} \|(\mathrm{Id} - P_n) x^\dagger\|_X \|w_n\|_Y,
\end{aligned} \tag{8.3}$$

wobei wir im letzten Schritt wieder (8.1) verwendet haben. Nach Annahme (8.2) ist nun der letzte Term beschränkt, während der vorletzte Term und damit die gesamte rechte Seite gegen Null geht. Zusammen mit Lemma 8.1 ergibt dies die Aussage. □

Damit die Ausgleichsprojektion einen Regularisierungsoperator definiert, müssen also die Unterräume X_n passend zu T gewählt werden. Bevor wir uns der dualen Ausgleichsprojektion zuwenden (die ohne eine solche Annahmen auskommt), betrachten wir noch den Spezialfall kompakter Operatoren.

Satz 8.3. *Erfüllt $K \in \mathcal{K}(X, Y)$ die Bedingung (8.2), so ist $x^\dagger \in \mathcal{R}(K^*)$.*

Beweis. Sei wieder $w_n := (K_n^\dagger)^* x_n$. Aus (8.2) folgt, dass $\{w_n\}_{n\in\mathbb{N}}$ beschränkt ist und daher eine schwach konvergente Teilfolge $w_k \rightharpoonup \bar{w} \in Y$ besitzt. Da K und damit auch K^* kompakt ist, folgt $K^* w_k \to K^* \bar{w}$. Nun ist aber wegen $(K_n^\dagger)^* = (K_n^*)^\dagger = (P_n K^*)^\dagger$

$$K^* w_k = P_k K^* w_k + (\mathrm{Id} - P_k) K^* w_k = x_k + (\mathrm{Id} - P_k) K^* w_k.$$

Grenzübergang auf beiden Seiten ergibt wegen Satz 8.2, der Beschränktheit von w_k, und $\|\mathrm{Id} - P_k\|_{\mathcal{L}(X,X)} \to 0$ nun $K^* \bar{w} = x^\dagger$, d. h. $x^\dagger \in \mathcal{R}(K^*)$. □

Die Bedingung (8.2) impliziert also bereits eine Quellbedingung. Es ist daher nicht verwunderlich, dass wir eine Abschätzung für die Konvergenz $x_n \to x^\dagger$ zeigen können.

Satz 8.4. *Erfüllt $K \in \mathcal{K}(X,Y)$ die Bedingung (8.2) und ist $y \in \mathcal{D}(K^\dagger)$, so existiert eine Konstante $C > 0$ mit*

$$\|x_n - x^\dagger\|_X \le C\|(\mathrm{Id} - P_n)K^*\|_{\mathcal{L}(Y,X)} \qquad \text{für alle } n \in \mathbb{N}.$$

Beweis. Wegen Satz 8.3 existiert ein $w \in Y$ mit $x^\dagger = K^* w$. Also gilt wegen (8.1)

$$\left(x_n - x^\dagger, x^\dagger\right)_X \le \|Kx_n - Kx^\dagger\|_Y \|w\|_Y \le \|K(P_n - \mathrm{Id})x^\dagger\|_Y \|w\|_Y.$$

Zusammen mit (8.3) und der Beschränktheit der $w_n := (K_n^\dagger)^* x_n$ folgt daraus

$$\begin{aligned}
\|x_n - x^\dagger\|_X^2 &= \left(x_n - x^\dagger, x_n\right)_X - \left(x_n - x^\dagger, x^\dagger\right)_X \\
&\le 2\|K(\mathrm{Id} - P_n)x^\dagger\|_Y \|w_n\|_Y + \|K(\mathrm{Id} - P_n)x^\dagger\|_Y \|w\|_Y \\
&\le C\|K(\mathrm{Id} - P_n)x^\dagger\|_Y = C\|K(\mathrm{Id} - P_n)(\mathrm{Id} - P_n)K^* w\|_Y \\
&\le C\|(\mathrm{Id} - P_n)K^*\|_{\mathcal{L}(Y,X)}^2 \|w\|_Y,
\end{aligned}$$

wobei wir im letzten Schritt verwendet haben, dass Projektionen selbstadjungiert sind und daher der Operator in der Norm die Form $\tilde{K}^*\tilde{K}$ für $\tilde{K} := (\mathrm{Id} - P_n)K^*$ hat. □

8.2 DUALE AUSGLEICHSPROJEKTION

Hier wird das Bild von T direkt diskretisiert. Wir betrachten also eine Folge $\{Y_n\}_{n\in\mathbb{N}}$ von geschachtelten Unterräumen

$$Y_1 \subset Y_2 \subset \cdots \subset \overline{\mathcal{R}(T)} = \mathcal{N}(T^*)^\perp \subset Y$$

mit $\dim Y_n = n$ und $\overline{\bigcup_{n\in\mathbb{N}} Y_n} = \mathcal{N}(T^*)^\perp$. Sei nun $Q_n := P_{Y_n}$ die orthogonale Projektion auf Y_n und $T_n := Q_n T \in \mathcal{L}(X,Y)$. Wieder ist $T_n^\dagger$ und damit auch $T_n^\dagger Q_n$ stetig, und wir können $x_n := T_n^\dagger Q_n y$, d. h. die Minimum-Norm-Lösung von $Q_n T x = Q_n y$, als Kandidaten für eine Regularisierung nehmen. Um zu zeigen, dass dadurch ein Regularisierungsoperator definiert wird, können wir folgende nützliche Charakterisierung verwenden.

Lemma 8.5. *Sei $y \in \mathcal{D}(T^\dagger)$. Dann ist $x_n = P_n x^\dagger$, wobei $P_n := P_{X_n}$ die orthogonale Projektion auf $X_n := T^* Y_n := \{T^* y : y \in Y_n\}$ bezeichnet.*

Beweis. Zuerst halten wir fest, dass wegen der Endlichdimensionalität von T_n gilt

$$\mathcal{R}(T_n^\dagger) = \mathcal{N}(T_n)^\perp = \mathcal{R}(T_n^*) = \mathcal{R}(T^* Q_n) = T^* Y_n = X_n$$

und damit $x_n \in X_n$ sowie $X_n^\perp = \mathcal{N}(T_n)$. Daraus folgt auch

$$T_n(\mathrm{Id} - P_n)x = 0 \qquad \text{für alle } x \in X,$$

d. h. $T_n P_n = T_n$. Weiterhin gilt wegen $Y_n \subset \mathcal{N}(T^*)^\perp = \overline{\mathcal{R}(T)}$ (und daher $\mathcal{R}(T)^\perp \subset \mathcal{N}(Q_n)$) sowie Lemma 3.4 (iv)

$$Q_n y = Q_n P_{\overline{\mathcal{R}(T)}} y = Q_n T T^\dagger y = Q_n T x^\dagger = T_n x^\dagger.$$

Zusammen erhalten wir für beliebige $x \in X$, dass gilt

$$\|T_n x - Q_n y\|_Y = \|T_n x - T_n x^\dagger\|_Y = \|T_n x - T_n P_n x^\dagger\|_Y = \|T_n(x - P_n x^\dagger)\|_Y.$$

Nun ist x_n definiert als die Minimum-Norm-Lösung von $T_n x = Q_n y$, d. h. als dasjenige $x \in \mathcal{N}(T_n)^\perp = X_n$, für das $\|T_n x - Q_n y\|_Y$ minimal wird – was für $x = P_n x^\dagger \in X_n$ der Fall ist. Da die Minimum-Norm-Lösung eindeutig ist, folgt $x_n = P_n x^\dagger$. □

Satz 8.6. *Sei $y \in \mathcal{D}(T^\dagger)$. Dann konvergiert $x_n \to x^\dagger$.*

Beweis. Aus den Eigenschaften von Y_n folgt $X_n \subset X_{n+1}$ und

$$\overline{\bigcup_{n\in\mathbb{N}} X_n} = \overline{\bigcup_{n\in\mathbb{N}} T^* Y_n} = \overline{T^* \bigcup_{n\in\mathbb{N}} Y_n} = \overline{T^* \mathcal{N}(T^*)^\perp} = \overline{\mathcal{R}(T^*)} = \mathcal{N}(T)^\perp.$$

Wegen $x^\dagger \in \mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$ konvergiert also $x_n \to x^\dagger$. □

Mit Hilfe einer Quellbedingung können wir wieder eine Fehlerabschätzung wie in Satz 8.4 zeigen.

Satz 8.7. *Für $T \in \mathcal{L}(X, Y)$ und $y \in \mathcal{D}(T^\dagger)$ mit $x^\dagger \in \mathcal{R}(T^*)$ existiert eine Konstante $C > 0$ mit*

$$\|x_n - x^\dagger\|_X \le C \|(\mathrm{Id} - P_n)T^*\|_{\mathcal{L}(Y,X)} \qquad \textit{für alle } n \in \mathbb{N}.$$

Beweis. Aus der Quellbedingung $x^\dagger = T^* w$ für ein $w \in Y$ und Lemma 8.5 folgt sofort

$$\|x_n - x^\dagger\|_X = \|P_n x^\dagger - x^\dagger\|_X = \|(\mathrm{Id} - P_n)T^* w\|_X \le \|(\mathrm{Id} - P_n)T^*\|_{\mathcal{L}(Y,X)} \|w\|_Y. \qquad \square$$

Auch die duale Ausgleichsprojektion definiert also einen Regularisierungsoperator. Nach Satz 4.5 existiert daher eine a priori-Parameterwahl, mit der die duale Ausgleichsprojektion ein konvergentes Regularisierungsverfahren bildet. Um diese Wahl zu charakterisieren, müssen wir die Norm von $T_n^\dagger$ abschätzen. Dafür können wir verwenden, dass T_n ein endlichdimensionales Bild hat und deshalb kompakt ist; es existiert also ein (endliches) singuläres System $\{(\mu_k, \tilde{u}_k, \tilde{v}_k\}_{k\in\{1,\dots,n\}}$. Insbesondere sei μ_n der betragskleinste Singulärwert von T_n.

Satz 8.8. *Sei $y \in \mathcal{D}(T^\dagger)$ sowie $y^\delta \in B_\delta(y)$ und $x_n^\delta = T_n^\dagger Q_n y$. Ist $n(\delta)$ so gewählt, dass gilt*

$$n(\delta) \to \infty, \qquad \frac{\delta}{\mu_{n(\delta)}} \to 0 \qquad \text{für } \delta \to 0,$$

dann konvergiert $x_{n(\delta)}^\delta \to x^\dagger$ für $\delta \to 0$.

Beweis. Wir gehen vor wie in Satz 4.6 und beginnen mit der üblichen Zerlegung

$$\|x_{n(\delta)}^\delta - x^\dagger\|_X \le \|x_{n(\delta)} - x^\dagger\|_X + \|x_{n(\delta)}^\delta - x_{n(\delta)}\|_X.$$

Der erste Term konvergiert nach Satz 8.7 gegen Null für $n \to \infty$.

Für den zweiten Term verwenden wir die Singulärwertzerlegung von T_n und (3.14) und erhalten für jedes $n \in \mathbb{N}$

$$\|T_n^\dagger y\|_X^2 = \sum_{k=1}^{n} \mu_k^{-2} |\,(y, \tilde{u}_k)_Y\,|^2 \le \mu_n^{-2} \|y\|_Y^2 \qquad \text{für alle } y \in Y,$$

mit Gleichheit für $y = \tilde{u}_n \in Y$, d. h. $\|T_n^\dagger\|_{\mathcal{L}(Y,X)} = \mu_n^{-1}$. Da Q_n eine orthogonale Projektion ist, gilt

$$\|x_n^\delta - x_n\|_X = \|T_n^\dagger Q_n (y^\delta - y)\|_X \le \|T_n^\dagger\|_{\mathcal{L}(Y,X)} \|y^\delta - y\|_Y \le \frac{\delta}{\mu_n}.$$

Die Behauptung folgt nun aus den Bedingungen an $n(\delta)$. □

Unter der Quellbedingung aus Satz 8.7 erhalten wir auf diesem Weg ähnlich wie in Satz 5.9 auch Konvergenzraten. (Analoge Resultate gelten unter der Bedingung (8.2) auch für die Ausgleichsprojektion.)

Wir können uns nun fragen, wie wir Y_n für festes n wählen müssen, um den Gesamtfehler zu minimieren, wofür wir nach Satz 8.8 insbesondere μ_n maximieren müssen. Diese Frage kann für kompakte Operatoren explizit mit Hilfe des *Courantschen Minimax-Prinzip*[2] für Eigenwerte beantwortet werden: Die (abfallend sortierten) Eigenwerte λ_n eines selbstadjungierten kompakten Operators $A \in \mathcal{K}(X, X)$ erfüllen

$$\begin{aligned} \lambda_n &= \min_{V \subset X} \max_{x \in V} \left\{ (Ax, x)_X : \|x\|_X = 1,\ \dim V^\perp = n - 1 \right\} \\ &= \max_{V \subset X} \min_{x \in V} \left\{ (Ax, x)_X : \|x\|_X = 1,\ \dim V = n \right\}. \end{aligned}$$

Satz 8.9. *Sei $K \in \mathcal{K}(X, Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n \in \mathbb{N}}$ und $Y_n \subset Y$ mit $\dim Y_n = n$. Dann ist $\mu_n \le \sigma_n$.*

[2] z. B. [Kaballo 2011, Satz 12.6]

Beweis. Da μ_n Singulärwert von K_n ist, ist μ_n^2 Eigenwert von $K_n K_n^* = Q_n K K^* Q_n$; analog ist σ_n^2 Eigenwert von KK^*. Sei für $k \in \mathbb{N}$ weiterhin $U_k := \operatorname{span}\{u_1, \dots, u_k\} \subset \overline{\mathcal{R}(K^*)}$. Wegen $\dim Y_n = n$ existiert ein $\bar{y} \in U_{n-1}^\perp \cap Y_n$ mit $\|\bar{y}\|_Y = 1$ (sonst müsste $U_{n-1}^\perp \subset Y_n^\perp$ sein, hat dafür aber zu kleine Kodimension). Aus dem Courantschen Minimax-Prinzip folgt dann

$$\begin{aligned}
\mu_n^2 &= \max_V \min_y \left\{ (Q_n K K^* Q_n y, y)_y : \|y\|_Y = 1,\ y \in V,\ \dim V = n \right\} \\
&= \min_y \left\{ (KK^* y, y)_Y : \|y\|_Y = 1,\ y \in Y_n \right\} \le (KK^* \bar{y}, \bar{y})_Y \\
&\le \max_y \left\{ (KK^* y, y)_Y : \|y\|_Y = 1,\ y \in U_{n-1}^\perp \right\} = \sigma_n^2,
\end{aligned}$$

denn das Maximum wird für $y = u_n \in U_{n-1}^\perp$ angenommen. □

Dabei wird Gleichheit für $Y_n = U_n$ angenommen, denn dann ist $\bar{y} = u_n$ der einzige Vektor, der im Beweis in Frage kommt. Dies entspricht aber genau der abgeschnittenen Singulärwertentwicklung aus Beispiel 5.2 (i). Tatsächlich ist die Wahl $Y_n = U_n$ auch optimal bezüglich des Verfahrensfehlers.

Satz 8.10. *Sei $K \in \mathcal{K}(X, Y)$ mit singulärem System $\{(\sigma_n, u_n, v_n)\}_{n\in\mathbb{N}}$ und $Y_n \subset Y$ mit $\dim Y_n = n$. Dann ist*

$$\|(\mathrm{Id} - P_n) K^*\|_{\mathcal{L}(Y,X)} \ge \sigma_{n+1},$$

wobei für $Y_n = U_n$ Gleichheit herrscht.

Beweis. Wir verwenden wieder das Courantsche Minimax-Prinzip, diesmal für σ_n^2 Eigenwert von K^*K. Mit $X_n := K^* Y_n$ und $P_n := P_{X_n}$ gilt dann

$$\begin{aligned}
\sigma_{n+1}^2 &= \min_V \max_x \left\{ (K^* K x, x)_X : \|x\|_X = 1,\ x \in V \subset X,\ \dim V^\perp = n \right\} \\
&\le \max_x \left\{ (K^* K x, x)_X : \|x\|_X = 1,\ x \in X_n^\perp \right\} \\
&= \max_x \left\{ (K^* K (\mathrm{Id} - P_n) x, (\mathrm{Id} - P_n) x)_X : \|x\|_X = 1 \right\} \\
&= \max_x \left\{ \|K(\mathrm{Id} - P_n) x\|_Y^2 : \|x\|_X = 1 \right\} \\
&= \|K(\mathrm{Id} - P_n)\|_{\mathcal{L}(X,Y)}^2 = \|(\mathrm{Id} - P_n) K^*\|_{\mathcal{L}(Y,X)}^2.
\end{aligned}$$

Ist $Y_n = U_n$, so ist $X_n = K^* U_n = \operatorname{span}\{v_1, \dots, v_n\}$, und für $x = v_{n+1} \in X_n^\perp$ wird das Minimum bzw. Maximum angenommen. □

Damit ist die bestmögliche Konvergenzrate (unter der Quellbedingung aus Satz 8.7) für die duale Ausgleichsprojektion

$$\|x_n^\delta - x^\dagger\|_X \le C \left(\sigma_{n+1} + \frac{\delta}{\sigma_n} \right),$$

und diese wird für die abgeschnittene Singulärwertzerlegung angenommen.

Ohne Kenntnis eines singulären Systems muss man in der Praxis n sehr klein wählen, um die Bedingung an μ_n garantieren zu können, und man erhält damit eine zu grobe Diskretisierung. Man kombiniert daher üblicherweise eine deutlich feinere Diskretisierung mit einer der bereits besprochenen Regularisierungen. Um dabei eine optimale Konvergenzrate zu erhalten und gleichzeitig unnötigen Rechenaufwand zu vermeiden, sollte dabei der Regularisierungsparameter in Abhängigkeit sowohl von δ als auch von n (bzw. n in Abhängigkeit von $\alpha(\delta)$) passend gewählt werden.

Teil III

STATISTISCHE INVERSE PROBLEME

9 FREQUENTISTISCHE INVERSE PROBLEME

Bislang hatten wir die gestörten Daten y^δ als ein beliebiges Element des Datenraums Y aufgefasst, von dem wir lediglich das Fehlerniveau $\delta = \|y^\delta - y\|_Y$ kennen. In *statistischen inversen Problemen* sind die gestörten Daten dagegen eine Zufallsvariable mit einer (hoffentlich) bekannten Verteilung. Statt dem schlimmstmöglichen Gesamtfehler über alle mit dem Fehlerniveau verträglichen Daten untersucht man daher den *durchschnittlichen* Gesamtfehler bezüglich dieser Verteilung über *alle* Daten. (Im Gegensatz dazu bezeichnet man den Ansatz aus den letzten Kapiteln als *deterministische inverse Probleme*.)

In der Statistik spricht man dabei auch von *Inferenz* und unterscheidet zwei Ansätze:

1. *frequentistische Inferenz*, bei der die exakte (Minimum-Norm-)Lösung $x^\dagger$ als fest gegeben vorausgesetzt wird, und
2. *Bayesianische Inferenz*, bei der auch die exakte (unbekannte) Lösung eine Zufallsvariable mit einer gewissen Verteilung ist.

Die Unterscheidung beruht in der Statistik auf tiefen philosophischen Differenzen über die Natur von Wahrscheinlichkeiten: Frequentisten betrachten diese als Grenzwert von relativen Häufigkeiten in Zufallsexperimenten, wenn die Anzahl der Wiederholungen gegen Unendlich geht (was für Messungen Sinn ergeben mag, nicht aber für die gesuchte exakte Lösung). Bayesianer fassen dagegen Wahrscheinlichkeiten auf als Quantifizierung von Unwissen (was dagegen auch – und gerade – für die exakte Lösung relevant ist). Unabhängig von der Philosophie führen die beiden Ansätze aber auch zu handfesten methodischen Unterschieden.

Wir betrachten in diesem Kapitel den ersten Ansatz, wobei wir in der Formulierung (wenn auch nicht der Notation) den Konventionen aus der Statistik folgen. Wie in den Kapiteln bisher untersuchen wir unendlichdimensionale Probleme; dies wird als *nichtparametrische Statistik* bezeichnet (im Unterschied zur parametrischen Statistik, in der ein durch endlich viele skalare Parameter beschriebenes – etwa lineares, polynomielles, oder exponentielles – Model betrachtet und dann diese Parameter geschätzt werden).

9.1 STATISTISCHES FEHLERMODELL UND SCHÄTZER

Wir müssen zuerst festlegen, was wir unter zufälligen Fehlern in unendlichdimensionalen Räumen verstehen. Seien X und Y wieder Hilberträume und $T \in \mathcal{L}(X, Y)$ linear und stetig. Wir nehmen an, für gegebenes $x \in X$ haben wir als Messung

$$y^\delta := Tx + \delta\xi, \tag{9.1}$$

wobei ξ eine zufällige Störung ist. Wir betrachten hier *Gaußsches weißes Rauschen*; solche Fehler tauchen in einer Vielzahl von praktischen Situationen auf und sind auch in gewisser Weise der einfachste Fall. (Alternativen sind z. B. *Laplace-* oder *Poisson-Rauschen*.) Die wesentliche Schwierigkeit ist, dass ξ – und damit y^δ – im Allgemeinen[1] *kein* Element von Y ist. Stattdessen ist es ein *stochastischer Prozess*, der ein $y \in Y$ linear und beschränkt abbildet auf eine reelle Zufallsvariable. (Der Fehler ξ ist also nur durch seine Auswirkung auf eine Messung von Tx definiert).

Wir benötigen dafür einige Begriffe aus der Wahrscheinlichkeitstheorie, für die wir auf die entsprechende Vorlesung verweisen. Ist Ω ein Ereignisraum, Σ eine Sigma-Algebra auf Ω, und $\mu : \Sigma \to [0, 1]$ ein Wahrscheinlichkeitsmaß, dann bezeichnet $L^2(\Omega; \mu)$ den Raum der bezüglich μ quadrat-integrierbaren Zufallsvariablen. Für Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ auf dem Wahrscheinlichkeitsraum (Ω, Σ, μ) sind dann Erwartungswert und Kovarianz definiert als

$$\mathbb{E}[X] := \int_\Omega X \, d\mu, \qquad \mathrm{Cov}[X, Y] := \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right].$$

Schreiben wir wie in der Funktionalanalysis üblich $\langle \xi, y \rangle_Y$ für die Anwendung von ξ auf ein $y \in Y$, führt das auf die folgende Definition.

Definition 9.1. Sei (Ω, Σ, μ) ein Wahrscheinlichkeitsraum. Wir nennen einen linearen und beschränkten Operator $\xi : Y \to L^2(\Omega; \mu)$ *weißes Rauschen* (mit Erwartungswert 0 und Kovarianz Id) auf Y, falls gilt

(i) $\mathbb{E}[\langle \xi, y \rangle_Y] = 0$ für alle $y \in Y$;

(ii) $\mathrm{Cov}[\langle \xi, y_1 \rangle_Y, \langle \xi, y_2 \rangle_Y] = (y_1, y_2)_Y$ für alle $y_1, y_2 \in Y$.

Ein weißes Rauschen ξ heißt *Gaußsch*, falls für alle $n \in \mathbb{N}$ und beliebige paarweise verschiedenen $y_1, \ldots, y_n \in Y$ die Variablen $\langle \xi, y_1 \rangle_Y, \ldots, \langle \xi, y_n \rangle_Y$ identisch und unabhängig normalverteilt sind.

Ist $K : X \to Y$ ein linearer kompakter Operator, dann können wir für $y \in Y$ insbesondere die Folge $\{u_n\}_{n \in \mathbb{N}} \subset Y$ singulärer Vektoren betrachten. In diesem Fall ist für alle $n \in \mathbb{N}$ die reelle Zufallsvariable

$$\xi_n := \langle \xi, u_n \rangle_Y$$

[1] und tatsächlich mit Wahrscheinlichkeit 1

identisch und unabhängig normalverteilt mit Mittelwert 0 und Standardabweichung 1, denn $\mathbb{E}[\xi_n] = 0$ und $\mathrm{Cov}[\xi_m, \xi_n] = (u_n, u_m)_Y = 1$ für $n = m$ und 0 sonst. Analog erhalten wir aus (9.1)

$$y_n^\delta := \langle y^\delta, u_n \rangle_Y = (Kx, u_n)_Y + \delta\xi_n = \sigma_n (x, v_n)_X + \delta\xi_n \qquad \text{für alle } n \in \mathbb{N}$$

bzw. äquivalent wegen $\sigma_n > 0$

$$x_n^\delta := \sigma_n^{-1} y_n^\delta = (x, v_n)_X + \frac{\delta}{\sigma_n}\xi_n \qquad \text{für alle } n \in \mathbb{N}. \tag{9.2}$$

Dies wird in der Statistik als *Folgenraummodell* bezeichnet; in dieser Formulierung besteht die Aufgabe der Inferenz nun daraus, $x_n := (x, v_n)_X$ aus den gegebenen x_n^δ zu schätzen. Dies ist wegen $\sigma_n \to 0$ für $n \to \infty$ natürlich nicht stabil für alle $n \in \mathbb{N}$ möglich. Für $\delta = 0$ erhalten wir aber aus (9.2) mit

$$\sum_{n\in\mathbb{N}} x_n v_n = \sum_{n\in\mathbb{N}} \sigma_n^{-1} y_n v_n = \sum_{n\in\mathbb{N}} \sigma_n^{-1} (y, u_n)_Y v_n = x^\dagger$$

die Minimum-Norm-Lösung wieder; vergleiche Satz 3.10. Es gilt also $x_n = (x, v_n)_X = (x^\dagger, v_n)_X$ für jede Lösung $x \in X$ von $Kx = y$.

Analog zu Kapitel 4 definiert man für $\delta > 0$ nun einen *linearen Schätzer*

$$x_\gamma^\delta := \sum_{n\in\mathbb{N}} \gamma_n x_n^\delta v_n$$

für eine geeignete Folge $\{\gamma_n\}_{n\in\mathbb{N}} \subset (0, \infty)$.

Offensichtlich definiert jeder Filter $\{\varphi_\alpha\}_{\alpha>0}$ zusammen mit einer Parameterwahlstrategie für α einen linearen Schätzer via

$$\gamma_n := \varphi_{\alpha(\delta, y^\delta)}(\sigma_n^2)\sigma_n^2, \qquad \text{für alle } n \in \mathbb{N}. \tag{9.3}$$

(Schätzer entsprechen also Regularisierungs*verfahren* in der deterministischen Theorie.) Insbesondere ist für

(i) die abgeschnittene Singulärwertzerlegung: $\gamma_n = \begin{cases} 1 & \text{falls } \sigma_n^2 \geq \alpha(\delta, y^\delta), \\ 0 & \text{sonst;} \end{cases}$

(ii) die Tikhonov-Regularisierung: $\gamma_n = \frac{\sigma_n^2}{\sigma_n^2 + \alpha(\delta, y^\delta)}$;

(iii) die Landweber-Regularisierung: $\gamma_n = 1 - (1 - \omega\sigma_n^2)^{1/\alpha(\delta, y^\delta)}$.

(Für die praktische Umsetzung verwendet man natürlich die Formulierung aus den entsprechenden Kapiteln anstelle des Folgenraummodells.)

Filter sind jedoch nicht die einzige Art, um lineare Schätzer zu definieren. Von besonderer Bedeutung für die Theorie (wenn auch nicht die Praxis) ist der *Pinsker-Schätzer*, definiert durch die Wahl

$$\gamma_n := \max\{0, 1 - \kappa_\delta a_n\}$$

für eine monoton wachsende Folge $\{a_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ und die Lösung $\kappa_\delta > 0$ von

$$\kappa_\delta \rho^2 - \delta^2 \sum_{n\in\mathbb{N}} \frac{a_n}{\sigma_n^2} \max\{0, 1 - \kappa_\delta a_n\} = 0 \tag{9.4}$$

für eine weitere Konstante $\rho > 0$. Der Parameter κ_δ entspricht dabei dem Regularisierungsparameter; der Pinsker-Schätzer beinhaltet also eine konkrete (a priori) Parameterwahlstrategie. Diese kann man auch expliziter charakterisieren als

$$\kappa_\delta = \frac{\sum_{n=1}^{N_\delta} \sigma_n^{-2} a_n}{\frac{\rho^2}{\delta^2} + \sum_{n=1}^{N_\delta} \sigma_n^{-2} a_n} \quad \text{für} \quad N_\delta := \max\left\{ N \in \mathbb{N} : \delta^2 \sum_{n=1}^{N} \sigma_n^2 a_n (a_N - a_n) \le \rho^2 \right\}. \tag{9.5}$$

9.2 RISIKO UND ORDNUNGSOPTIMALITÄT

Da der Fehler und damit die Daten zufällig sind, ist auch der darauf beruhende Schätzer eine Zufallsvariable. Wir definieren anstelle des schlimmstmöglichen Gesamtfehlers (4.3) für einen (nicht notwendigerweise linearen) Schätzer $x(y^\delta)$ das *Risiko*

$$R(x(y^\delta), x^\dagger) := \mathbb{E}[\|x(y^\delta) - x^\dagger\|_X^2], \tag{9.6}$$

wobei der Erwartungswert über alle Störungen ξ gebildet wird. (Beachte, dass das Risiko auf dem *quadrierten* Gesamtfehler beruht.) Genau wie bei den deterministischen Fehlern basiert die Analysis nun auf der fundamentalen Zerlegung des Gesamtfehlers.

Satz 9.2. *Sei x_γ^δ ein linearer Schätzer für das Folgenraummodell (9.2). Dann gilt*

$$R(x_\gamma^\delta, x^\dagger) = \sum_{n\in\mathbb{N}} \left((1-\gamma_n)^2 x_n^2 + \delta^2 \frac{\gamma_n^2}{\sigma_n^2} \right).$$

Beweis. Zunächst folgt aus der Definition von x_γ^δ und der Charakterisierung von $x^\dagger$ in Satz 3.10

$$\|x_\gamma^\delta - x^\dagger\|_X^2 = \sum_{n\in\mathbb{N}} (\gamma_n x_n^\delta - x_n)^2.$$

Wir schieben nun wieder die produktive Null $\gamma_n x_n - \gamma_n x_n$ ein, verwenden (9.2), multiplizieren aus, und erhalten wegen der Linearität des Erwartungswerts

$$\begin{aligned}\mathbb{E}[\|x_\gamma^\delta - x^\dagger\|_X^2] &= \sum_{n\in\mathbb{N}} \mathbb{E}\left[\left((\gamma_n x_n^\delta - \gamma_n x_n) + (\gamma_n x_n - x_n)\right)^2\right] \\ &= \sum_{n\in\mathbb{N}} \mathbb{E}\left[\left(\gamma_n \frac{\delta}{\sigma_n}\xi_n + (\gamma_n - 1)x_n\right)^2\right] \\ &= \sum_{n\in\mathbb{N}} \left((1-\gamma_n)^2 x_n^2 + 2(1-\gamma_n)x_n \cdot \gamma_n \frac{\delta}{\sigma_n^2}\,\mathbb{E}[\xi_n] + \delta^2 \frac{\gamma_n^2}{\sigma_n^2}\,\mathbb{E}[\xi_n^2]\right).\end{aligned}$$

Da die ξ_n nach Annahme normalverteilt sind mit Mittelwert 0 und Kovarianz 1, gilt $\mathbb{E}[\xi_n] = 0$ und damit $\mathbb{E}[\xi_n^2] = \mathrm{Cov}[\xi_n, \xi_n] = 1$, woraus die Aussage folgt. □

Der erste Term in der Summe ist dabei genau der Verfahrensfehler und wird mit *Bias* bezeichnet; der zweite Term entspricht dem fortgepflanzten Datenfehler und wird mit *Varianz* bezeichnet. Wie bei der spektralen Regularisierung besteht nun die Aufgabe darin, die Gewichte γ_n (entsprechend der Parameterwahl für α) so zu wählen, dass die beiden Terme optimal ausbalanciert werden. Ist der Schätzer etwa wie in (9.3) durch einen Filter $\{\varphi_\alpha\}_{\alpha>0}$ definiert und setzen wir wieder $r_\alpha(\lambda) = 1 - \lambda\varphi_\alpha(\lambda)$, so gilt

$$R(x_\gamma^\delta, x^\dagger) = \sum_{n\in\mathbb{N}} \left(r_{\alpha(\delta,y^\delta)}(\sigma_n)^2 x_n^2 + \delta^2 \varphi_{\alpha(\delta,y^\delta)}(\sigma_n^2)^2 \sigma_n^2\right)$$

(wobei im Gegensatz zum schlimmstmöglichen Fehler hier sogar Gleichheit gilt). Der erste (rein deterministische) Term kann dabei exakt wie in der spektralen Regularisierung behandelt werden. Der zweite Term muss aber anders abgeschätzt werden: wir können nun nicht wie in Lemma 5.4 einfach $\sigma_n^2 \varphi_\alpha(\sigma_n^2) \le C_\varphi^2$ und dann den verbleibenden Term mit der Besselschen Ungleichung durch $\|y - y^\delta\|_Y^2 \le \delta^2$ abschätzen – das ist genau der Preis, den wir für den stochastischen Fehler $\xi \notin Y$ zahlen müssen.

Wie für deterministische inverse Probleme kann man nun versuchen, das Risiko unter einer Quellbedingung abzuschätzen. Wie in Lemma 4.13 zeigt man leicht, dass für das Folgenraummodell (9.2) die Quellbedingung $x \in X_{\nu,\rho}$ für $\nu, \rho > 0$ erfüllt ist genau dann, wenn gilt

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2\nu} x_n^2 \le \rho^2. \tag{9.7}$$

(Natürlich kann man – auch für deterministische Probleme – allgemeinere Gewichte $a_n \to \infty$ anstelle von $\sigma_n^{-\nu}$ betrachten; man spricht dann von *Koeffizienten-Ellipsoiden*.) Die Konsequenz des unterschiedlichen Varianz-Terms in Satz 9.2 ist dabei, dass für stochastische Probleme die Singulärwerte von K eine Rolle für die Konvergenzrate spielen. Wir zeigen dies beispielhaft für die abgeschnittene Singulärwertzerlegung, wobei wir uns auf moderat schlechtgestellte Probleme beschränken.

Satz 9.3. *Sei $K \in \mathcal{K}(X,Y)$ mit Singulärwerten*

$$c_\mu n^{-\mu} \le \sigma_n \le C_\mu n^{-\mu} \qquad \text{für alle } n \in \mathbb{N}$$

mit $\mu > 0$ und $C_\mu > c_\mu > 0$ sowie $x^\dagger \in X_{\nu,\rho}$ für $\nu, \rho > 0$. Ist x_γ^δ ein linearer Schätzer mit

$$\gamma_n = \begin{cases} 1 & \text{für } n \le N(\delta), \\ 0 & \text{sonst,} \end{cases} \qquad \text{mit} \qquad c_N \delta^{\frac{-2}{2\mu(\nu+1)+1}} \le N(\delta) \le C_N \delta^{\frac{-2}{2\mu(\nu+1)+1}} \tag{9.8}$$

für $C_N > c_N > 0$, so existiert ein $C > 0$ mit

$$R(x_\gamma^\delta, x^\dagger) \le C\delta^{\frac{4\mu\nu}{2\mu(\nu+1)+1}} \qquad \text{für } \delta \to 0. \tag{9.9}$$

Beweis. Aus Satz 9.2 folgt zunächst für diese Wahl von γ_n mit $N := N(\delta)$ wegen der Endlichkeit der zweiten Summe

$$R(x_\gamma^\delta, x^\dagger) = \sum_{n=N+1}^{\infty} x_n^2 + \delta^2 \sum_{n=1}^{N} \sigma_n^{-2}.$$

Für den ersten Term verwenden wir die Quellbedingung und erhalten

$$\sum_{n=N+1}^{\infty} x_n^2 = \sum_{n=N+1}^{\infty} \sigma_n^{-2\nu} x_n^2 \sigma_n^{2\nu} \le \sigma_N^{2\nu} \rho^2 \le C_\mu N^{-2\mu\nu} \rho^2$$

nach Annahme an die Singulärwerte.

Für den zweiten Term folgt ebenfalls aus der Annahme an die Singulärwerte[2]

$$\delta^2 \sum_{n=1}^{N} \sigma_n^{-2} \le \delta^2 c_\mu \sum_{n=1}^{N} n^{2\mu} \le \delta^2 c_\mu N \cdot N^{2\mu} = c_\mu \delta^2 N^{2\mu+1}.$$

Einsetzen der Parameterwahlregel (9.8) in beide Abschätzungen und Vereinfachen ergibt jetzt (9.9). □

Im Vergleich zu der nach Satz 5.9 und Beispiel 5.10 für die abgeschnittene Singulärwertzerlegung ordnungsoptimalen „deterministischen“ Rate $\delta^{\frac{2\nu}{\nu+1}} = \delta^{\frac{4\mu\nu}{2\mu(\nu+1)}}$ für den quadrierten schlimmstmöglichen Fehler ist die „statistische“ Rate (9.9) also schlechter und hängt noch dazu von dem Verhalten der Singulärwerte ab. Da die Rate – und die Parameterwahl – also noch mehr Kenntnis von in der Praxis unbekannten Koeffizienten benötigt, ist man in der Statistik in der Regel nicht so sehr an den konkreten Raten für einen Schätzers interessiert. Stattdessen definiert man „Ordnungsoptimalität“ direkt über das kleinstmögliche Risiko

[2] Dies ist für feste $N \in \mathbb{N}$ zwar nur eine grobe obere Schranke; die *Faulhabersche Formel* zeigt aber, dass $\sum_{n=1}^{N} n^r = \frac{1}{r+1} N^{r+1} + O(N^r)$ für $N \to \infty$ ist für alle $r > -1$. Diese Formel kann man zum Beispiel über die Euler–Maclaurinsche Formel (und für $r \in \mathbb{N}$ direkt per Induktion und binomischer Formel) beweisen.

über eine gewisse Klasse von Schätzern. Man nennt einen Schätzer $x(y^\delta)$ einen *linearen Minimax-Schätzer* für $X_{\nu,\rho}$, falls gilt

$$\sup_{x^\dagger \in X_{\nu,\rho}} R(x(y^\delta), x^\dagger) = \inf_\gamma \sup_{x^\dagger \in X_{\nu,\rho}} R(x_\gamma^\delta, x^\dagger),$$

wobei das Infimum über alle Folgen $\{\gamma_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ gebildet wird. Gilt lediglich

$$\sup_{x^\dagger \in X_{\nu,\rho}} R(x(y^\delta), x^\dagger) \leq C \inf_\gamma \sup_{x^\dagger \in X_{\nu,\rho}} R(x_\gamma^\delta, x^\dagger)$$

für eine Konstante $C > 0$, sagt man, $x(y^\delta)$ habe *(lineare) Minimax-Rate.*

Den linearen Minimax-Schätzer kann man sogar explizit angeben.

Satz 9.4. *Der Pinsker-Schätzer ist ein linearer Minimax-Schätzer. Konkret gilt für alle $\nu, \rho > 0$*

$$\inf_\gamma \sup_{x^\dagger \in X_{\nu,\rho}} R(x_\gamma^\delta, x^\dagger) = \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2} \max\{0, 1 - \kappa_\delta \sigma_n^{-\nu}\}$$

mit κ_δ Lösung von (9.4) für ρ und $a_n := \sigma_n^{-\nu}$, und das Infimum bzw. Supremum wird angenommen für

$$\bar\gamma_n = \max\{0, 1 - \kappa_\delta \sigma_n^{-\nu}\},$$
$$\bar x_n^2 = \frac{\delta^2}{\kappa_\delta} \sigma_n^{-2+\nu} \max\{0, 1 - \kappa_\delta \sigma_n^{-\nu}\}.$$

Beweis. Nach Satz 9.2 gilt für alle $x^\dagger$ und jeden linearen Schätzer

$$R(x_\gamma^\delta, x^\dagger) = \sum_{n\in\mathbb{N}} \left((1-\gamma_n)^2 x_n^2 + \delta^2 \frac{\gamma_n^2}{\sigma_n^2} \right).$$

Wir bilden nun zuerst das Infimum über alle γ_n für festes $x^\dagger$, was wegen der Nicht-Negativität für jeden Summanden einzeln möglich ist. Jeder Summand hat nun die Form $f(t) = a(1-t)^2 + bt^2$ mit $a, b > 0$, wofür wir durch elementare Kurvendiskussion (Setzen der Ableitung gleich Null, Auflösen, und Einsetzen) das Minimum $\min_t f(t) = \frac{ab}{a+b}$ finden. Es gilt also

$$\inf_\gamma R(x_\gamma^\delta, x^\dagger) = \sum_{n\in\mathbb{N}} \frac{\delta^2 \sigma_n^{-2} x_n^2}{x_n^2 + \delta^2 \sigma_n^{-2}}.$$

Speziell für $\bar x_n^2 = \frac{\delta^2}{\kappa_\delta} \sigma_n^{-2+\nu} \bar\gamma_n$ und $\bar\gamma_n = 1 - \kappa_\delta \sigma_n^{-\nu} > 0$ gilt weiter

$$\frac{\bar x_n^2}{\bar x_n^2 + \delta^2 \sigma_n^{-2}} = \frac{\bar\gamma_n}{\bar\gamma_n + \kappa_\delta \sigma_n^{-\nu}} = 1 - \kappa_\delta \sigma_n^{-\nu} = \bar\gamma_n.$$

(Für $\bar\gamma_n = 0$ ist auch $\bar x_n = 0$ und damit die Gleichung trivialerweise erfüllt.) Schließlich folgt aus (9.4)

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2\nu}\bar x_n^2 = \frac{\delta^2}{\kappa_\delta}\sum_{n\in\mathbb{N}} \sigma_n^{-\nu-2}\max\{0, 1-\kappa_\delta\sigma_n^{-\nu}\} = \rho^2$$

und damit $\bar x := \sum_{n\in\mathbb{N}} \bar x_n v_n \in X_{\nu,\rho}$. Da stets gilt $\inf\sup \geq \sup\inf$, erhalten wir daraus

$$\begin{aligned}(9.10)\qquad \inf_\gamma \sup_{x^\dagger\in X_{\nu,\rho}} R(x_\gamma^\delta, x^\dagger) &\geq \sup_{x^\dagger\in X_{\nu,\rho}} \inf_\gamma R(x_\gamma^\delta, x^\dagger) = \sup_{x^\dagger\in X_{\nu,\rho}} \sum_{n\in\mathbb{N}} \frac{\delta^2\sigma_n^{-2}x_n^2}{x_n^2+\delta^2\sigma_n^{-2}} \\ &\geq \sum_{n\in\mathbb{N}} \frac{\delta^2\sigma_n^{-2}\bar x_n^2}{\bar x_n^2+\delta^2\sigma_n^{-2}} = \delta^2\sum_{n\in\mathbb{N}} \sigma_n^{-2}\bar\gamma_n.\end{aligned}$$

Für die umgekehrte Richtung verwenden wir, dass für alle $x^\dagger \in X_{\nu,\rho}$ gilt

$$\sum_{n\in\mathbb{N}}(1-\bar\gamma_n)^2 x_n^2 \leq \sup_{n\in\mathbb{N}}\{(1-\bar\gamma_n)^2\sigma_n^{2\nu}\}\sum_{n\in\mathbb{N}}\sigma_n^{-2\nu}x_n^2 \leq \sup_{n\in\mathbb{N}}\{(1-\bar\gamma_n)^2\sigma_n^{2\nu}\}\rho^2.$$

Wieder machen wir die Fallunterscheidung

$$(1-\bar\gamma_n)^2\sigma_n^{2\nu} = \begin{cases}\sigma_n^{2\nu} & \text{falls } 1-\kappa_\delta\sigma_n^{-\nu}\leq 0,\\ \kappa_\delta^2 & \text{falls } 1-\kappa_\delta\sigma_n^{-\nu} > 0.\end{cases}$$

In beiden Fällen können wir daher durch κ_δ^2 abschätzen. Aus (9.4) folgt dann

$$\begin{aligned}(9.11)\qquad \inf_\gamma \sup_{x^\dagger\in X_{\nu,\rho}} R(x_\gamma^\delta, x^\dagger) &\leq \sup_{x^\dagger\in X_{\nu,\rho}} \sum_{n\in\mathbb{N}}\left((1-\bar\gamma_n)^2x_n^2 + \delta^2\frac{\bar\gamma_n^2}{\sigma_n^2}\right) \leq \kappa_\delta^2\rho^2 + \delta^2\sum_{n\in\mathbb{N}}\sigma_n^{-2}\bar\gamma_n^2 \\ &= \delta^2\sum_{n\in\mathbb{N}}\sigma_n^{-2}(\kappa_\delta\sigma_n^{-\nu}\bar\gamma_n + \bar\gamma_n^2) = \delta^2\sum_{n\in\mathbb{N}}\sigma_n^{-2}\bar\gamma_n,\end{aligned}$$

wobei man die letzte Gleichung einmal mehr durch Fallunterscheidung für $\bar\gamma_n > 0$ erhält. □

Tatsächlich bieten die KKT-Bedingungen für den linearen Minimax-Schätzer (als konvexes Optimierungsproblem für γ unter der Nebenbedingung $x^\dagger \in X_{\nu,\rho}$) eine mögliche Herleitung der Gleichung (9.4) sowie der Charakterisierung ihrer Lösung. Mit deutlich mehr Aufwand kann man auch zeigen, dass der Pinsker-Schätzer für $\delta \to 0$ die Minimax-Rate über *alle* Schätzer hat, siehe [Belitser & Levit 1995].

Der Pinsker-Schätzer bietet nun eine Messlatte für andere lineare Schätzer. Zum Beispiel folgt aus der Tatsache, dass die Singulärwerte eine monotone Nullfolge bilden, dass $\bar\gamma_n = 0$ ist für alle $n > N$ mit $\sigma_N^\nu < \kappa_\delta \leq \sigma_{N+1}^\nu$, ähnlich wie bei der abgeschnittenen Singulärwertzerlegung. Tatsächlich basiert die Optimalität des Pinsker-Schätzers im Wesentlichen auf der Optimalität dieser Parameterwahl.

Folgerung 9.5. *Ist $x_{\hat{\gamma}}^{\delta}$ ein linearer Schätzer mit*

$$\hat{\gamma}_n = \begin{cases} 1 & \text{für } n \le N(\delta), \\ 0 & \text{sonst,} \end{cases} \qquad \text{mit} \qquad \sigma_n^{\nu} < 2\kappa_\delta \quad \text{für alle } n > N(\delta) \tag{9.12}$$

mit κ_δ Lösung von (9.4) für ρ und $a_n := \sigma_n^{-\nu}$, so gilt

$$\sup_{x^\dagger \in X_{\nu,\rho}} R(x_{\hat{\gamma}}^{\delta}, x^\dagger) \le 4 \inf_{\gamma} \sup_{x^\dagger \in X_{\nu,\rho}} R(x_{\gamma}^{\delta}, x^\dagger).$$

Beweis. Im Beweis von Satz 9.3 haben wir gezeigt, dass gilt

$$\sup_{x^\dagger \in X_{\nu,\rho}} R(x_{\hat{\gamma}}^{\delta}, x^\dagger) \le \sigma_{N(\delta)}^{2\nu}\rho^2 + \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2}.$$

Nun ist nach Definition $\bar{\gamma}_n = 1 - \kappa_\delta \sigma_n^{-\nu} \ge \frac{1}{2}$ für alle $n \le N(\delta)$, und daher gilt

$$\sigma_{N(\delta)}^{2\nu}\rho^2 + \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2} \le (2\kappa_\delta)^2\rho^2 + \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2}(2\bar{\gamma}_n)^2.$$

Aus Satz 9.4 folgt weiter, dass alle Ungleichungen in (9.10) und (9.11) mit Gleichheit gelten, und daher gilt

$$\kappa_\delta^2\rho^2 + \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2}\bar{\gamma}_n^2 = \inf_{\gamma} \sup_{x^\dagger \in X_{\nu,\rho}} R(x_{\gamma}^{\delta}, x^\dagger),$$

und zusammen erhalten wir die Aussage. □

Der nächste Schritt besteht nun darin, unter der Annahme $\sigma_n \approx n^{-\mu}$ aus dem Ansatz $\sigma_{N(\delta)}^{\nu} \approx \kappa_\delta$ mit Hilfe der Gleichung (9.4) eine Abschätzung für $N(\delta) \approx \delta^{\frac{-2}{2\mu(\nu+1)+1}}$ herzuleiten und damit zu zeigen, dass die abgeschnittene Singulärwertzerlegung auch mit der a priori Wahl (9.8) die Minimax-Rate hat. Wir sparen uns das hier und begnügen uns mit einer Plausibilitätsschätzung (sämtliche Konstanten und Terme niedrigerer Ordnung werden einfach weggelassen):

$$N(\delta)^{-\mu\nu} \approx \sigma_{N(\delta)}^{\nu} \approx \kappa_\delta \approx \delta^2 \sum_{n\in\mathbb{N}} \sigma_n^{-2-\nu}\bar{\gamma}_n \approx \delta^2 \sum_{n=1}^{N(\delta)} n^{2\mu+\mu\nu} \approx \delta^2 N(\delta)^{2\mu+\mu\nu+1},$$

wobei wir im letzten Schritt wieder wie im Beweis von Satz 9.3 abgeschätzt haben. Auflösen nach $N(\delta)$ ergibt dann die gewünschte Abschätzung. Allgemein kann man zeigen, dass regularisierende Filter zusammen mit der entsprechenden a priori-Parameterwahl Schätzer definieren, die (bis zu ihrer Qualifikation) die lineare Minimax-Rate annehmen; siehe [Bissantz u. a. 2007].

10 BAYESIANISCHE INVERSE PROBLEME

Zum Abschluß sei hier ein kurzer Ausblick auf den Bayesianischen Zugang zu statistischen inversen Problemen gegeben. Wie im letzten Kapitel erwähnt, beruht dieser darauf, dass keine exakte (Minimum-Norm-)Lösung $x^\dagger$ im Hintergrund festgehalten wird, sondern diese ebenfalls durch eine Wahrscheinlichkeitsverteilung beschrieben wird – entsprechend dem Grundsatz „Zufall ist Unwissen“, und die exakte Lösung ist eben nicht bekannt. Eine alternative Interpretierung ist als das Ersetzen der harten Quellbedingung $x^\dagger \in X_\nu$ durch eine Verteilung für die gesuchte Minimum-Norm-Lösung, ebenso wie im letzten Kapitel die harte Bedingung $y^\delta \in B_\delta(Kx^\dagger)$ durch eine Verteilung für das Rauschen ersetzt wurde.

Wieder benötigen wir ein paar Begriffe aus der Wahrscheinlichkeitstheorie. Sei (Ω, Σ, μ) ein Wahrscheinlichkeitsraum. Aus technischen Gründen, auf die wir hier nur am Rande eingehen, nehmen wir hier $\Omega \subset \mathbb{R}^N$ für $N \in \mathbb{N}$ an; entsprechend sei Σ als die Borel-Algebra und μ als das Lebesgue-Maß festgelegt. Ist $X : \Omega \to \mathbb{R}$ eine Zufallsvariable, dann wird dadurch ein neues Wahrscheinlichkeitsmaß auf $\mathbb{R}$ versehen mit der Borel-Algebra festgelegt durch

$$\mu_X(A) := \mu(X^{-1}(A)) \qquad \text{für alle Borel-meßbaren } A \subset \mathbb{R},$$

d. h. die Wahrscheinlichkeit, dass $X \in A$ liegt. Man nennt dann X *verteilt nach* μ_X und schreibt $X \sim \mu_X$. Existiert weiter ein $\rho_X \in L^1(\Omega; \mu)$ mit

$$\mu_X(A) := \int_A \rho_X(x)\, d\mu \qquad \text{für alle } \Sigma\text{-meßbaren } A \subset \Omega,$$

so nennt man ρ_X *Dichte* von μ_X (bezüglich μ). (Für $\Omega \subset X$ mit X ein unendlichdimensionaler Hilbertraum existiert im Allgemeinen kein dem Lebesgue-Maß entsprechendes Referenzmaß, bezüglich dem man Dichten allgemein definieren kann.)

Seien nun $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen. Die gemeinsame Verteilung ist dann

$$\mu_{X,Y}(A \times B) := \mu(X^{-1}(A) \cap Y^{-1}(B)) \qquad \text{für alle Borel-meßbaren } A, B \subset \mathbb{R}.$$

Gilt

$$\mu_{X,Y}(A \times B) = \mu_X(A)\mu_Y(B) \qquad \text{für alle Borel-meßbaren } A, B \subset \mathbb{R},$$

so heißen X und X *unabhängig*; in diesem Fall hat $\mu_{X,Y}$ die Dichte $\rho_{X,Y}(x, y) = \rho_X(x)\rho_Y(y)$. Auf diese Weise kann man Zufallsvariablen mit Werten in $\mathbb{R}^N$ definieren. Haben wir

umgekehrt eine gemeinsame Verteilung $\mu_{X,Y}$ mit Dichte $\rho_{X,Y}$ gegeben, so erfüllt die *Randverteilung*

$$\mu_X(A) = \mu_{X,Y}(A \times \Omega) \qquad \text{für alle Borel-meßbaren } A \subset \mathbb{R}.$$

Entsprechend definiert man die *bedingte Wahrscheinlichkeit*

$$\mu_{X|Y\in B}(A) := \frac{\mu_{X,Y}(A \times B)}{\mu_Y(B)} \qquad \text{für alle Borel-meßbaren } A, B \subset \mathbb{R}.$$

Diese Definition kann man auch auf Punktmengen der Form $B = \{y_0\}$ (die ja Lebesgue-Nullmengen sind) sinnvoll erweitern (mit Hilfe regulärer bedingter Wahrscheinlichkeiten). Wir schreiben in diesem Fall $\mu_{X|y_0}$ für $y_0 \in Y$ als die *bedingte Wahrscheinlichkeitsverteilung für X gegeben y_0*.

Der *Satz von Bayes* erlaubt nun, diese bedingte Wahrscheinlichkeitsverteilung durch die gemeinsame Verteilung und die Randverteilung für X zu charakterisieren (wobei der Beweis bei der richtigen Definition aller Begriffe eine reine Anwendung des Satzes von Fubini ist). Wir geben her direkt eine Variante für Dichten an, die wir in Folge anwenden werden.

Satz 10.1. *Seien $X : \Omega \to \mathbb{R}^N$ und $Y : \Omega \to \mathbb{R}^M$ Zufallsvariablen. Haben μ_X die Dichte ρ_X, μ_Y die Dichte ρ_Y, und $\mu_{Y|x}$ für μ_X-fast alle $x \in \mathbb{R}^N$ die Dichte $\rho_{Y|x}$, dann hat $\mu_{X|y}$ für μ_Y-fast alle $y \in \mathbb{R}^M$ die Dichte*

$$\rho_{X|y}(x) = \frac{\rho_{Y|x}(y)\rho_X(x)}{\rho_Y(y)} \qquad \textit{für } \mu_X\textit{-fast alle } x \in \mathbb{R}^M,$$

wobei gilt

$$\rho_Y(y) = \int_{\mathbb{R}^N} \rho_{Y|x}(y)\, d\mu_X = \int_{\mathbb{R}^N} \rho_{Y|x}(y)\rho_X(x)\, dx.$$

Diesen Satz wenden wir nun auf inverse Probleme der Form $Tx = y$ für $T \in \mathcal{L}(\mathbb{R}^N \to \mathbb{R}^M)$ an. Speziell wählen wir

- für die Unbekannte $x^\dagger \in \mathbb{R}^N$ eine *a priori-Verteilung* μ_X basierend auf etwaigem Vorwissen (english „prior“);
- für die Messung $y^\delta \in \mathbb{R}^M$ eine *Plausibilität* $\mu_{Y|x}$, die angibt, wie wahrscheinlich eine Messung y^δ für angenommene exakte Daten $y := Tx \in \mathbb{R}^N$ wäre (english „likelihood“).

Die *a posteriori-Verteilung* $\mu_{X|y^\delta}$ (englisch „posterior“) sagt uns dann, wie konsistent ein beliebiges $x \in \mathbb{R}^N$ sowohl mit unserem Vorwissen als auch mit einer Messung y^δ ist, und quantifiziert damit sämtliche verbleibende Unsicherheit nach der Messung. Die *Evidenz* $\rho_Y(y^\delta)$ ist dabei eine Normierungskonstante, die dafür sorgt, dass die a posteriori-Verteilung tatsächlich eine Wahrscheinlichkeitsverteilung ist.

Die wesentlichen Fragen sind nun

1. Wie sind a priori-Verteilung und Plausibilität problemspezifisch zu wählen?
2. Welche praktischen Informationen können aus der a posteriori-Verteilung gewonnen werden?

Wir skizzieren in Folge jeweils die einfachsten Ansätze, wobei hervorzuheben ist, dass diese die Stärke des Bayesianischen Zugangs nicht ansatzweise repräsentieren.

Wir beginnen mit der Modellierung, wo wir – sowohl der Einfachheit halber als auch um den Zusammenhang mit den früheren Kapiteln herzustellen – für die Unbekannte und die Messung jeweils eine Gaußsche Normalverteilung annehmen. Wir erinnern, dass eine Zufallsvariable $X : \Omega \to \mathbb{R}$ normalverteilt ist mit Mittelwert t_0 und Varianz σ^2 genau dann, wenn sie die Dichte

$$\rho_X(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(t-t_0)^2} \qquad \text{für alle } t \in \mathbb{R}$$

hat; wir schreiben dann $X \sim N(t_0, \sigma^2)$. Haben wir N unabhängig und identisch normalverteilte Variablen X_i gegeben, so ist $X := (X_1, \dots, X_N)^T : \Omega \to \mathbb{R}^N$ ebenfalls normalverteilt mit Dichte

$$\rho_X(x) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-x_0)^2} = \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2\sigma^2}\|x-x_0\|_2^2} \qquad \text{für alle } x \in \mathbb{R}^N.$$

Allgemein sagen wir, dass für $x_0 \in \mathbb{R}^N$ und $C \in \mathbb{R}^{N\times N}$ symmetrisch und positiv definit die Zufallsvariable $X : \Omega \to \mathbb{R}^N$ normalverteilt ist mit Erwartungswert x_0 und *Kovarianzmatrix* C und schreiben $X \sim N(x_0, C)$, falls ihre Dichte gegeben ist durch

$$\rho_X(x) = \frac{1}{\sqrt{(2\pi)^N \det C}} e^{-\frac{1}{2}\left(C^{-1}(x-x_0),(x-x_0)\right)_2} \qquad \text{für alle } x \in \mathbb{R}^N. \tag{10.1}$$

(Ist K injektiv, so kann man $x \sim N(0, C)$ für $C = |K|^{-2\nu} = (K^*K)^{-\nu}$ als Bayesianisches Äquivalent der Quellbedingung $x \in X_\nu$ auffassen.)

Wir machen nun für a priori-Verteilung und Plausibilität die Annahmen

- $\mu_X = N(0, \sigma^2 \operatorname{Id})$ für ein $\sigma > 0$;
- $\mu_{Y|x} = N(Tx, \delta^2 \operatorname{Id})$ für ein $\delta > 0$.

(Letzteres entspricht genau der Annahme in Kapitel 9; da wir hier nur endlichdimensionale Probleme betrachten, müssen wir nicht mit stochastischen Prozessen arbeiten. In der Praxis sollten diese Verteilungen natürlich aus einer sorgfältigen statistischen Modellierung des Problems kommen.)

Nach Satz 10.1 und den Rechenregeln für Exponentialfunktionen hat nun die a posteriori-Verteilung $\mu_{X|y^\delta}$ für eine gegebene Messung $y^\delta \in \mathbb{R}^M$ die Dichte

$$\rho_{X|y^\delta}(x) = \frac{e^{-\frac{1}{2\delta^2}\|Tx-y^\delta\|_2^2-\frac{1}{2\sigma^2}\|x\|_2^2}}{\int_{\mathbb{R}^N} e^{-\frac{1}{2\delta^2}\|Tx-y^\delta\|_2^2-\frac{1}{2\sigma^2}\|x\|_2^2}\,dx} \qquad \text{für alle } x \in \mathbb{R}^N. \tag{10.2}$$

Diese Dichte kann nun verwendet werden, um Aussagen darüber zu treffen, welche Punkte $x \in \mathbb{R}$ besonders wahrscheinlich zu dieser Messung geführt haben; man spricht dabei von *Punktschätzern*. Der am häufigsten verwendete Punktschätzer ist der *Maximum-a-posteriori-* oder *MAP-Schätzer*, definiert als der Punkt, der (in einem geeigneten Sinn) die maximale Wahrscheinlichkeit unter der a posteriori-Verteilung hat. In unserem Fall ist dies der (globale) Maximierer der Dichte, d. h.

$$\begin{aligned} x_{\text{MAP}} &:= \arg\max_{x\in\mathbb{R}^N} \rho_{X|y^\delta}(x) = \arg\min_{x\in\mathbb{R}^N} -\log \rho_{X|y^\delta}(x) \\ &= \arg\min_{x\in\mathbb{R}^N} \frac{1}{2\delta^2}\|Tx-y^\delta\|_2^2 + \frac{1}{2\sigma^2}\|x\|_2^2. \end{aligned}$$

In anderen Worten x_{MAP} entspricht – für diese Wahl von a priori-Verteilung und Plausibilität![1] – nach Satz 6.5 genau der Tikhonov-Regularisierung mit $\alpha = \frac{\delta^2}{\sigma^2}$, die nach Lemma 6.3 auch geschrieben werden kann als

$$x_{\text{MAP}} = \left(T^*T + \frac{\delta^2}{\sigma^2}\,\text{Id}\right)^{-1} T^* y^\delta. \tag{10.3}$$

Eine Alternative ist der *bedingte Erwartungswert* („conditional mean“)

$$x_{\text{CM}} := \mathbb{E}[X|Y = y^\delta] = \int_{\mathbb{R}^N} x\,d\mu_{X|y^\delta} = \int_{\mathbb{R}^N} x\rho_{X|y^\delta}(x)\,dx.$$

Um diesen Schätzer für unseren konkreten Ansatz zu berechnen, schreiben wir die a posteriori-Dichte (10.2) etwas um, indem wir ausmultiplizieren, (10.3) verwenden, und quadratisch ergänzen. Wir erhalten dadurch

$$\begin{aligned} \frac{1}{2\delta^2}\|Tx-y^\delta\|_2^2 + \frac{1}{2\sigma^2}\|x\|_2^2 &= \frac{1}{2\delta^2}\left((T^*T+\tfrac{\delta^2}{\sigma^2}\,\text{Id})x, x\right)_2 - \frac{1}{\delta^2}\left(T^*y^\delta, x\right)_2 + \frac{1}{2\delta^2}\|y^\delta\|_2^2 \\ &= \frac{1}{2\delta^2}\left((T^*T+\tfrac{\delta^2}{\sigma^2}\,\text{Id})(x-x_{\text{MAP}}), x\right)_2 \\ &\quad - \frac{1}{2\delta^2}\left((T^*T+\tfrac{\delta^2}{\sigma^2}\,\text{Id})x, x_{\text{MAP}}\right)_2 + \frac{1}{2\delta^2}\|y^\delta\|_2^2 \\ &= \frac{1}{2\delta^2}\left((T^*T+\tfrac{\delta^2}{\sigma^2}\,\text{Id})(x-x_{\text{MAP}}), x-x_{\text{MAP}}\right)_2 \\ &\quad - \frac{1}{2\delta^2}\left((T^*T+\tfrac{\delta^2}{\sigma^2}\,\text{Id})x_{\text{MAP}}, x_{\text{MAP}}\right)_2 + \frac{1}{2\delta^2}\|y^\delta\|_2^2. \end{aligned}$$

[1] Definiert man sich umgekehrt a priori-Verteilung und Plausibilität bewusst so, dass am Ende als MAP-Schätzer genau die Tikhonov-Regularisierung herauskommt, ist das ein *Bayesianisches Verbrechen*.

Die letzten beiden Terme sind dabei konstant bezüglich x. Außerdem ist $T^*T + \frac{\delta^2}{\sigma^2}\,\mathrm{Id}$ selbstadjungiert und positiv definit und hat daher eine ebenfalls selbstadjungierte und positiv definite Inverse. Bis auf eine Konstante hat daher (10.2) genau die Form (10.1), d. h. die a posteriori-Verteilung ist ebenfalls normalverteilt mit

$$\mu_{X|y^\delta} \sim N\left(x_{\mathrm{MAP}}, \delta^2 (T^*T + \tfrac{\delta^2}{\sigma^2}\,\mathrm{Id})^{-1}\right). \tag{10.4}$$

Insbesondere folgt $x_{\mathrm{CM}} = x_{\mathrm{MAP}}$. Ist jedoch die a priori-Verteilung oder die Plausibilität nicht normalverteilt, oder ist das inverse Problem nichtlinear, so ist der bedingte Erwartungswert im Allgemeinen nicht gleich dem MAP-Schätzer. In diesem Fall hat man auch keine geschlossene Darstellung von x_{CM} und muss eine numerische Approximation verwenden. Nach Satz 10.1 gilt

$$x_{\mathrm{CM}} = \int_{\mathbb{R}^N} x \rho_{X|y^\delta}(x)\,dx = \frac{\int_{\mathbb{R}^N} x \rho_{Y|x}(y^\delta)\rho_X(x)\,dx}{\int_{\mathbb{R}^N} \rho_{Y|x}(y^\delta)\rho_X(x)\,dx} = \frac{\int_{\mathbb{R}^N} x \rho_{Y|x}(y^\delta)\,d\mu_X}{\int_{\mathbb{R}^N} \rho_{Y|x}(y^\delta)\,d\mu_X}$$

wegen der Definition der Dichte. Dies sind jeweils hochdimensionale Integrale, so dass eine klassische Quadraturformel nicht in Frage kommt. Stattdessen verwendet man *Monte Carlo-Integration*: Sind $x_1, \dots, x_n$ unabhängig und identisch entsprechend μ_X verteilte Stichproben (welche zumindest für die Normalverteilung einfach zu erzeugen sind), so ist

$$\hat{x}_{\mathrm{CM}} := \frac{\frac{1}{n}\sum_{i=1}^n x_i \rho_{Y|x_i}(y^\delta)}{\frac{1}{n}\sum_{i=1}^n \rho_{Y|x_i}(y^\delta)} = \frac{\sum_{i=1}^n x_i e^{-\frac{1}{2\delta^2}\|Tx_i - y^\delta\|_2^2}}{\sum_{i=1}^n e^{-\frac{1}{2\delta^2}\|Tx_i - y^\delta\|_2^2}}$$

eine Näherung von x_{CM}, die nach dem Gesetz der großen Zahlen mit der Rate $O(1/\sqrt{n})$ gegen x_{CM} konvergiert. (Eine Alternative sind *Metropolis–Hastings Markov-Ketten Monte Carlo-Verfahren*, die direkt eine Folge von Stichproben nach der a posteriori-Verteilung erzeugen.)

Monte Carlo-Verfahren können auch verwendet werden, um *Bereichsschätzer* aus der a posteriori-Verteilung zu gewinnen; solche Schätzer können die verbleibende Unsicherheit quantifizieren (man spricht daher auch von „uncertainty quantification“). Ein Beispiel sind *Glaubwürdigkeitsmengen* (englisch „credible sets“); dies sind für gegebenes $\alpha \in (0,1)$ Mengen $C_\alpha \subset \mathbb{R}^N$, für die gilt

$$\int_{C_\alpha} = \mu_{X|y^\delta}(C_\alpha) = \rho_{X|y^\delta}(x)\,dx = 1 - \alpha, \tag{10.5}$$

d. h. C_α enthält $1-\alpha$ (z. B. 0.95) der Masse der a posteriori-Verteilung. Je größer die Menge, desto unsicherer ist man sich also über die wahre Lösung (die aber nicht in C_α liegen muss!) Diese Mengen sind für gegebenes α nicht eindeutig; man kann sich zum Beispiel auf *Mengen der höchsten a posteriori-Dichte*

$$C_\alpha^* = \left\{x \in \mathbb{R}^N : -\log \rho_{X|y^\delta}(x) \le \eta_\alpha\right\}$$

beschränken, so dass es ausreicht, ein $\eta_\alpha \in (0,\infty)$ zu finden, so dass (10.5) erfüllt ist.

LITERATUR

H. W. Alt (2012), *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*, 6. Aufl., Springer-Verlag, DOI: 10.1007/978-3-642-22261-0.

A. B. Bakushinskiĭ (1985), Remarks on choosing a regularization parameter using the quasioptimality and ratio criterion, *USSR Comput. Math. Math. Phys.* 24(4), 181–182, DOI: 10.1016/0041-5553(84)90253-2.

F. Bauer & M. A. Lukas (2011), Comparing parameter choice methods for regularization of ill-posed problems, *Mathematics and Computers in Simulation* 81(9), 1795–1841, DOI: 10.1016/j.matcom.2011.01.016.

E. N. Belitser & B. Y. Levit (1995), On minimax filtering over ellipsoids, *Math. Methods Stat.* 4(3), 259–273.

N. Bissantz, T. Hohage, A. Munk & F. Ruymgaart (2007), Convergence rates of general regularization methods for statistical inverse problems and applications, *SIAM J. Numer. Anal.* 45(6), 2610–2636, DOI: 10.1137/060651884.

M. Burger (2007), Inverse Problems, Vorlesungsskript, Institut für Numerische und Angewandte Mathematik, Universität Münster, URL: http://wwwmath.uni-muenster.de/num/Vorlesungen/IP_WS07/skript.pdf.

L. Cavalier (2011), Inverse problems in statistics, in: *Inverse Problems and High-Dimensional Estimation*, Bd. 203, Lect. Notes Stat. Proc. Springer, Heidelberg, 3–96, DOI: 10.1007/978-3-642-19989-9_1.

C. Clason (2019), *Einführung in die Funktionalanalysis*, Mathematik Kompakt, Birkhäuser, Basel, DOI: 10.1007/978-3-030-24876-5,

N. Du (2008), Finite-dimensional approximation settings for infinite-dimensional Moore–Penrose inverses, *SIAM J. Numer. Anal.* 46(3), 1454–1482, DOI: 10.1137/060661120.

H. W. Engl, M. Hanke & A. Neubauer (1996), *Regularization of Inverse Problems*, Bd. 375, Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, DOI: 10.1007/978-94-009-1740-8.

T. Hohage (2000), Regularization of exponentially ill-posed problems, *Numer. Funct. Anal. Optim.* 21(3-4), 439–464, DOI: 10.1080/01630560008816965.

T. Hohage (2002), Inverse Problems, Vorlesungsskript, Institut für Numerische und Angewandte Mathematik, Universität Göttingen.

K. Ito & B. Jin (2014), *Inverse Problems: Tikhonov Theory and Algorithms*, Bd. 22, Series on Applied Mathematics, World Scientific, Singapore, DOI: 10.1142/9789814596206_0001.

W. Kaballo (2011), *Grundkurs Funktionalanalysis*, Heidelberg: Spektrum Akademischer Verlag, DOI: 10.1007/978-3-8274-2721-2.

H. Kekkonnen (2019), Bayesian Inverse Problems, Vorlesungsskript, University of Cambridge, URL: http://www.damtp.cam.ac.uk/research/cia/files/teaching/Inverse_Problems_Lent_2018/19y_01m_16d_LectureNotes.pdf.

S. Kindermann (2011), Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems, *Electron. Trans. Numer. Anal.* 38, 233–257, URL: http://etna.mcs.kent.edu/volumes/2011-2020/vol38/abstract.php?vol=38&pages=233-257.

S. Kindermann (2016), Projection methods for ill-posed problems revisited, *Comput. Methods Appl. Math.* 16(2), 257–276, DOI: 10.1515/cmam-2015-0036.

A. Kirsch (2011), *An Introduction to the Mathematical Theory of Inverse Problems*, 2. Aufl., Springer, New York, DOI: 10.1007/978-1-4419-8474-6.

L. Landweber (1951), An iteration formula for Fredholm integral equations of the first kind, *Amer. J. Math.* 73, 615–624, DOI: 10.2307/2372313.

A. K. Louis (1989), *Inverse und schlecht gestellte Probleme*, Teubner Studienbücher Mathematik, B. G. Teubner, Stuttgart, DOI: 10.1007/978-3-322-84808-6.

P. Mathé & B. Hofmann (2008), How general are general source conditions?, *Inverse Problems* 24(1), 015009, 5, DOI: 10.1088/0266-5611/24/1/015009.

A. Rieder (2003), *Keine Probleme mit inversen Problemen*, Eine Einführung in ihre stabile Lösung, Friedr. Vieweg & Sohn, Braunschweig, DOI: 10.1007/978-3-322-80234-7.

A. N. Tikhonov (1963a), On the regularization of ill-posed problems, *Dokl. Akad. Nauk SSSR* 153, 49–52.

A. N. Tikhonov (1963b), On the solution of ill-posed problems and the method of regularization, *Dokl. Akad. Nauk SSSR* 151, 501–504.

B. von Harrach (2014), Regularisierung Inverser Probleme, Vorlesungsskript, Fachbereich Mathematik, Universität Stuttgart, URL: http://www.mathematik.uni-stuttgart.de/~harrach/lehre/Regularisierung.pdf.

D. Werner (2011), *Funktionalanalysis*, 7. Aufl., Springer-Verlag, Berlin, DOI: 10.1007/978-3-642-21017-4.