Vortrag zum Artikel
"An iterative thresholding algorithm for linear inverse problems with a sparsity constraint "
von I. Daubechies, M. Defrise, C. De Mol

Thomas Bonesky

24. Juni 2005

Inhaltsverzeichnis

Kapitel 1

Einleitung

Dieser Vortrag behandelt ein Lösungsverfahren für inverse Probleme der Form

$$Kf = h \tag{1.1}$$

mit einem linearen, beschränkten, kompakten Operator $K : H \to H'$ zwischen separablen Hilberträumen. Wobei h die bekannte, beziehungsweise gemessene und f die gesuchte Größe ist.
In Anwendungen besteht die rechte Seite der Gleichung in der Regel aus fehlerbehafteten Messdaten. Messfehler werden oft durch einen additiven Fehlerterm (hier e) modelliert.

$$Kf + e = h + e = g \tag{1.2}$$

Ein klassisches Verfahren zur Lösung eines solchen inversen Problems ist die Tikhonov-Regularisierung, die auf der Minimierung des Tikhonov-Funktionals

$$T_\mu(f) = ||Kf - g||^2 + \mu||f||^2 \tag{1.3}$$

beruht, wobei der zweite Summand den sogenannten Strafterm darstellt. μ ist ein geeignet zu wählender Regularisierungsparameter. Für dieses Verfahren kann man die Lösung recht einfach bestimmen. Die Tikhonov-Regularisierung hat jedoch einen entscheidenden Nachteil. Sie sorgt dafür, dass die berechneten Lösungen stark geglättet sind. Um schärfere Konturen in den Lösungen zu erreichen kann man ein neues Funktional definieren und den Strafterm wie folgt ersetzen:

$$\Phi_{w,p}(f) = ||Kf - g||^2 + \sum_{\gamma \in \Gamma} w_\gamma |\langle f, \varphi_\gamma \rangle|^p \tag{1.4}$$

dabei ist $w = (w_\gamma)_{\gamma\in\Gamma}$ eine strikt positive Folge, $(\varphi_\gamma)_{\gamma\in\Gamma}$ eine Orthonormalbasis des Hilbertraumes H und $1 \le p \le 2$. Dieser veränderte Strafterm

dient dazu Normen anderer Funktionenräume darzustellen, um somit schärfere Konturen der Lösung auflösen zu können.
Setzt man $w_\gamma = \mu \; \forall \gamma \in \Gamma$ und $p = 2$, so erhält man aus diesem Ansatz die Tikhonov-Regularisierung als Spezialfall.
Mit diesem verallgemeinerten Ansatz kann nun das Problem der geglätteten Lösungen umgangen werden, allerdings tritt ein neues Problem auf. Während es bei der Tikhonov-Regularisierung recht einfach ist einen Minimierer des Funktionals zu finden, stellt sich dies nun weitaus komplizierter dar.
Die Minimierung des Funktionals $\Phi_{w,p}$ führt auf folgende Variationsgleichungen:

$$\forall \gamma \in \Gamma : \langle K^*Kf, \varphi_\gamma \rangle - \langle K^*g, \varphi_\gamma \rangle + \frac{w_\gamma p}{2} |\langle f, \varphi_\gamma \rangle|^{p-1} \mathrm{sign}(\langle f, \varphi_\gamma \rangle) = 0 \quad (1.5)$$

Die Schwierigkeiten bei der Lösung dieses Systems liegen darin begründet, dass die Gleichungen zum einen nichtlinear bezüglich der $\langle f, \varphi_\gamma \rangle$ sind und alle Gleichungen über den Operator K^*K gekoppelt sind. Um das System lösen zu können muss man sich eines der beiden Probleme entledigen. Dies führt im Weiteren auf die Einführung von Ersatzfunktionalen. Das daraus resultierende Lösungsverfahren ist Inhalt des folgenden Abschnitts.

Kapitel 2

Einführung des Minimierungsalgorithmus

Um die Kopplung der Gleichungen in (1.5) zu umgehen werden Ersatzfunktionale eingeführt. Dazu wird zunächst das Funktional

$$\Theta(f;a) := C||f-a||^2 - ||Kf-Ka||^2, \tag{2.1}$$

mit einem beliebig fest gewählten Element $a \in H$ und einer Konstanten $C > ||K^*K||$, definiert. Nachfolgend wird angenommen, dass $||K|| < 1$ gilt. Somit kann $C = 1$ gesetzt werden. Es kann gezeigt werden, dass das Funktional $\Theta(\,\cdot\,;a)$ strikt konvex ist. Daher ergibt sich auch für das wie folgt definierte Ersatzfunktional strikte Konvexität.

$$\Phi_{w,p}^{SUR}(f;a) := \Phi_{w,p}(f) - ||Kf-Ka||^2 + ||f-a||^2 \tag{2.2}$$

Nach einigen Umformungen und mit der Kurzschreibweise $v_\gamma := \langle v, \varphi_\gamma \rangle$ resultiert für das Ersatzfunktional die Form:

$$\Phi_{w,p}^{SUR}(f;a) = \sum_{\gamma\in\Gamma} \left[f_\gamma^2 - 2f_\gamma(a+K^*g-K^*Ka)_\gamma + w_\gamma|f_\gamma|^p\right] + ||g||^2 + ||a||^2 - ||Ka||^2 \tag{2.3}$$

Nun wird durch

$$f^0 \;\text{ beliebig};\quad f^n = \operatorname{argmin}\left(\Phi_{w,p}^{SUR}(f;f^{n-1})\right);\quad n = 1,2,\ldots \tag{2.4}$$

ein Algorithmus definiert, von dem erhofft wird, dass er einen Minimierer des Ausgangsfunktionals $\Phi_{w,p}$ liefert. Das die so definierte Folge $(f_n)_{n\in\mathbb{N}}$ jedenfalls schwach gegen einen Minimierer des Ausgangsfunktionals konvergiert wird im nächsten Abschnitt bewiesen. Doch zunächst sollen einige wichtige Eigenschaften dieses Ansatzes erläutert werden.

Die Minimierung des Ersatzfunktionals führt auf entkoppelte Variationsgleichungen. Für $p > 1$ ergibt sich

$$\forall \gamma \in \Gamma : f_\gamma + \frac{pw_\gamma}{2}\text{sign}(f_\gamma)|f_\gamma|^{p-1} = (a_\gamma + [K^*(g - Ka)]_\gamma). \tag{2.5}$$

Zur Vorbereitung der Lösung dieses Systems wird zunächst ein Lemma bewiesen.

LEMMA 2.1
Für $w > 0$ und $p > 1$ ist die Funktion $F_{w,p} : \mathbb{R} \to \mathbb{R}$ mit

$$F_{w,p}(x) = x + \frac{wp}{2}\text{sign}(x)|x|^{p-1}$$

bijektiv.

Beweis:
Für $x \geq 0$ ergibt sich $F_{w,p}$ als $F_{w,p}(x) = x + \frac{wp}{2}x^{p-1}$. Diese Funktion ist stetig und streng monoton wachsend als Zusammensetzung solcher Funktionen. Außerdem gilt offenbar $x \to \infty \implies F_{w,p}(x) \to \infty$.
Für $x \leq 0$ ergibt sich $F_{w,p}$ als $F_{w,p}(x) = x - \frac{wp}{2}(-x)^{p-1}$. Auch diese Funktion ist stetig und streng monoton wachsend nach Bauart und es gilt $x \to -\infty \implies F_{w,p}(x) \to -\infty$.
Es gilt jeweils $F_{w,p}(0) = 0$, womit die Aussage bewiesen ist. □

Nach Lemma 2.1 hat die Funktion $F_{w,p}$ also auf ganz $\mathbb{R}$ eine eindeutige Inverse. Sei also

$$S_{w,p} = (F_{w,p})^{-1}, \tag{2.6}$$

dann gilt für die gesuchten Minimierer

$$f_\gamma = S_{w_\gamma,p}(a_\gamma + [K^*(g - Ka)]_\gamma). \tag{2.7}$$

Für den Fall $p = 1$ kann man ebenfalls eine Funktion $S_{w,1}$ finden, über die sich die Minimierer bestimmen lassen.

Zusammenfassend ergibt sich der folgende Satz.

SATZ 2.2
Sei $K : H \to H'$ ein linearer, beschränkter, kompakter Operator zwischen separablen Hilberträumen mit $||K^*K|| < 1$ und sei $g \in H'$. Sei $(\varphi_\gamma)_{\gamma\in\Gamma}$ eine Orthonormalbasis von H und $(w_\gamma)_{\gamma\in\Gamma}$ eine strikt positive Folge. Dann hat das mit einem beliebigen $p \geq 1$ und $a \in H$ als

$$\Phi^{SUR}_{w,p}(f; a) = \Phi_{w,p}(f) - ||Kf - Ka||^2 + ||f - a||^2 \tag{2.8}$$

definierte Funktional einen eindeutigen Minimierer.
Der Minimierer wird gegeben durch

$$f = S_{w,p}(a + K^*(g - Ka)), \tag{2.9}$$

wobei die Operatoren $S_{w,p}$ definiert sind durch

$$S_{w,p}(h) = \sum_{\gamma \in \Gamma} S_{w_\gamma,p}(h_\gamma)\varphi_\gamma. \tag{2.10}$$

Darüber hinaus gilt für alle $h \in H$

$$\Phi_{w,p}^{SUR}(f + h; a) \geq \Phi_{w,p}^{SUR}(f; a) + ||h||^2. \tag{2.11}$$

zum Beweis:
Die Eindeutigkeit des Minimierers folgt aus der strikten Konvexität des Funktionals. Die Darstellung des Minimieres folgt mit Lemma 2.1. Die letzte Aussage lässt sich durch Einsetzen der expliziten Ausdrücke für $S_{w_\gamma,p}$ und einige Fallunterscheidungen zeigen.

Aus diesem Satz folgt unmittelbar:

KOROLLAR 2.3
Mit den Voraussetzungen von Satz 2.2 und dem in (2.4) beschriebenen Algorithmus ergibt sich

$$f^n = S_{w,p}(f^{n-1} + K^*(g - Kf^{n-1})). \tag{2.12}$$

Kapitel 3

Schwache Konvergenz des Minimierungsverfahrens

In diesem Abschnitt geht es darum, die schwache Konvergenz, der in Korollar 2.3 beschriebenen Folge der Minimierer der Ersatzfunktionale, gegen einen Minimierer von $\Phi_{w,p}$ zu zeigen.
Nach Einführung der Notation

$$Tf := S_{w,p}(f + K^*(g - Kf)) \tag{3.1}$$

wird die schwache Konvergenz der Folge $f^n = T^n f^0$ mit Hilfe des Fixpunktsatzes von Opial bewiesen.

SATZ 3.1 (Opial)
Sei H ein Hilbertraum. Erfüllt die Abbildung $A : H \to H$ die folgenden Bedingungen

- A ist nicht-expansiv: $\forall v, v' \in H$, $||Av - Av'|| \leq ||v - v'||$,
- A ist asymptotisch regulär: $\forall v \in H$, $||A^{n+1}v - A^n v|| \stackrel{n\to\infty}{\longrightarrow} 0$,
- die Menge $\mathcal{F}$ der Fixpunkte von A ist nicht leer,

dann gilt, $\forall v \in H$ konvergiert die Folge $(A^n v)_{n\in\mathbb{N}}$ schwach gegen einen Fixpunkt in $\mathcal{F}$.

In diesem Fall kann beispielsweise der klassische Fixpunktsatz von Banach nicht verwendet werden, da er an die Abbildung A statt der Bedingung nicht-expansiv, die Bedingung kontrahierend stellt, d.h.

$$\forall v, v' \in H \text{ gilt } ||Av - Av'|| \leq q||v - v'|| \tag{3.2}$$

mit $q < 1$.
Für den hier betrachteten Operator T kann jedoch nur gezeigt werden, dass er nicht-expansiv ist. Mit dem Satz 3.1 folgt daher im Gegensatz zum Satz von Banach nur die schwache Konvergenz.
Um Satz 3.1 auf den Operator T anwenden zu können, müssen zunächst die Voraussetzungen des Satzes überprüft werden. Beispielhaft wird an dieser Stelle die asymptotische Regularität bewiesen. Zu diesem Zweck werden einige Lemmas benötigt. Vorab werden noch nachstehende Vorbemerkungen gemacht.

Vorbemerkungen:
Ist $K : H \to H'$ ein linearer, beschränkter und kompakter Operator zwischen separablen Hilberträumen, so ist der Operator K^*K darüber hinaus selbstadjungiert. Es lassen sich die Eigenwerte von K^*K $\lambda_0 \geq \lambda_1 \geq \ldots$ und orthonormierte Eigenvektoren v_n, $n \geq 0$ bestimmen, so dass mit Singulärwerten $\sigma_n := +\sqrt{\lambda_n}$ und Vektoren $u_n := \sigma_n^{-1} K v_n$ gilt:

$$\begin{aligned} K v_n &= \sigma_n u_n \\ K^* u_n &= \sigma_n v_n \end{aligned} \tag{3.3}$$

Darüber hinaus stellt die Menge $\{v_n\}$ ein vollständiges Orthonormalsystem für $\mathrm{Kern}(K)^\perp$ dar.

DEFINITION UND SATZ 3.3
Die Menge von Tripel

$$\{v_n, u_n; \sigma_n\}_{n \geq 0} \tag{3.4}$$

heißt singuläres System für K.

Der Operator K besitzt die Darstellung

$$K \cdot = \sum_n \sigma_n \langle \cdot, v_n \rangle u_n. \tag{3.5}$$

Das oben definierte vollständige Orthonormalsystem für $\mathrm{Kern}(K)^\perp$ lässt sich zu einer Basis von H ergänzen. Dabei werden die "zugehörigen" Singulärwerte gleich Null gesetzt.

Nun zum Nachweis der asymptotischen Regularität.

LEMMA 3.2
Sowohl $(\Phi_{w,p}(f^n))_{n\in\mathbb{N}}$ als auch $(\Phi_{w,p}^{SUR}(f^{n+1}; f^n))_{n\in\mathbb{N}}$ sind monoton-fallende Folgen.

Beweis:
Nach Aussagen der Funktionalanalysis und mit Hilfe der Vorbemerkungen lässt sich der Operator

$$L\cdot := \sqrt{I - K^*K}\cdot = \sum_n \sqrt{(1-\sigma_n^2)}\langle \cdot, v_n\rangle v_n \tag{3.6}$$

definieren.
Mit dieser Definition gilt für ein $h \in H$

$$\begin{aligned} ||h||^2 - ||Kh||^2 &= \sum_n |\langle h, v_n\rangle|^2 - \sum_n \sigma_n^2 |\langle h, v_n\rangle|^2 \\ &= \sum_n (1-\sigma_n^2)|\langle h, v_n\rangle|^2 \\ &= \sum_n \left(\sqrt{(1-\sigma_n^2)}\right)^2 |\langle h, v_n\rangle|^2 \\ &= ||Lh||^2 \end{aligned} \tag{3.7}$$

Nun ergibt sich, da f^{n+1} das Funktional $\Phi_{w,p}^{SUR}(f; f^n)$ minimiert

$$\begin{aligned} \Phi_{w,p}^{SUR}(f^{n+2}; f^{n+1}) &\leq \Phi_{w,p}^{SUR}(f^{n+1}; f^{n+1}) \\ &= \Phi_{w,p}(f^{n+1}) \\ &\leq \Phi_{w,p}(f^{n+1}) + ||L(f^{n+1} - f^n)||^2 \\ &= \Phi_{w,p}(f^{n+1}) + ||f^{n+1} - f^n||^2 - ||Kf^{n+1} - Kf^n||^2 \\ &= \Phi_{w,p}^{SUR}(f^{n+1}; f^n) \\ &\leq \Phi_{w,p}^{SUR}(f^n; f^n) \\ &= \Phi_{w,p}(f^n) \end{aligned} \tag{3.8}$$

womit die Behauptungen folgen. □

LEMMA 3.3
Die Reihe $\sum_{n=0}^{\infty} ||f^{n+1} - f^n||^2$ ist konvergent.

Beweis:
Es gilt $\lambda_0 = \sigma_0^2 = ||K^*K|| < 1$ nach Voraussetzung, wobei λ_0 der größte Eigenwert von K^*K ist. Damit folgt

$$\begin{aligned} ||Lf||^2 &= ||\sum_n \sqrt{(1-\sigma_n^2)}\langle f, v_n\rangle v_n||^2 \\ &= \sum_n \underbrace{(1-\sigma_n^2)}_{\geq(1-\sigma_0^2)=:A} |\langle f, v_n\rangle|^2 \\ &\geq A\sum_n |\langle f, v_n\rangle|^2 \\ &= A||f||^2. \end{aligned} \tag{3.9}$$

Mit Lemma 3.2 gilt dann $\forall N \in \mathbb{N}$ die Abschätzung

$$\begin{aligned} \sum_{n=0}^{N} ||f^{n+1} - f^n||^2 &\leq \frac{1}{A}\sum_{n=0}^{N} ||L(f^{n+1} - f^n)||^2 \\ &\leq \underbrace{\frac{1}{A}\sum_{n=0}^{N}[\Phi_{w,p}(f^n) - \Phi_{w,p}(f^{n+1})]}_{\text{Teleskopsumme}} \\ &= \frac{1}{A}[\Phi_{w,p}(f^0) - \Phi_{w,p}(f^{N+1})] \\ &\leq \frac{1}{A}\,\Phi_{w,p}(f^0) \end{aligned} \tag{3.10}$$

Nun ist gezeigt, dass die Folge $\sum_{n=0}^{N} ||f^{n+1} - f^n||^2$ gleichmäßig beschränkt ist, daraus ergibt sich die Behauptung. $\square$

Aus Lemma 3.3 folgt jetzt unmittelbar

LEMMA 3.4
Die Abbildung $T\cdot = S_{w,p}(\cdot + K^*(g - K\cdot))$ ist asymptotisch regulär, d.h.

$$||T^{n+1}f^0 - T^n f^0|| = ||f^{n+1} - f^n|| \overset{n\to\infty}{\longrightarrow} 0 \tag{3.11}$$

Auch die weiteren Voraussetzungen zur Anwendung des Satzes von Opial können gezeigt werden, so dass der folgende Satz unmittelbar aus dem Satz von Opial folgt.

SATZ 3.5
Die Folge $f^n = T^n f^0$, $n \in \mathbb{N}$ konvergiert schwach gegen einen Fixpunkt von T.

Schließlich ist noch zu zeigen

SATZ 3.6
Jeder Fixpunkt von T minimiert das Funktional $\Phi_{w,p}$.

Beweis:
Sei $f^\star$ ein Fixpunkt von T, dann gilt nach Definition der Ersatzfunktionale und mit Satz 2.2 $\forall h \in H$

$$\begin{aligned} \Phi_{w,p}(f^\star + h) &= \Phi_{w,p}^{SUR}(f^\star + h; f^\star) - ||h||^2 + ||Kh||^2 \\ &\geq \Phi_{w,p}^{SUR}(f^\star; f^\star) + ||Kh||^2 \\ &= \Phi_{w,p}(f^\star) + ||Kh||^2. \end{aligned} \tag{3.12}$$

Damit ist die Behauptung gezeigt.□

Zusammenfassend ergibt sich also der

SATZ 3.7
Sei $K : H \to H'$ ein beschränkter, linearer, kompakter Operator zwischen separablen Hilberträumen mit $||K|| < 1$. Wählt man weiter ein $p \in [1, 2]$ und eine strikt positive Folge $(w_\gamma)_{\gamma\in\Gamma}$, dann konvergiert die Folge $f^n = T^n f^0$, $n > 0$, mit beliebigem $f^0 \in H$, schwach gegen einen Minimierer des Funktionals $\Phi_{w,p}$. Gilt darüber hinaus $\mathrm{Kern}(K) = \{0\}$ oder $p > 1$ so ist der Minimierer eindeutig.

zum Beweis:
Bis auf die Eindeutigkeit sind bereits alle Aussagen gezeigt.
Die Eindeutigkeit folgt, da $\Phi_{w,p}$ in den beschriebenen Fällen strikt konvex ist.

Kapitel 4

Schlussbemerkung

Über die schwache Konvergenz des Minimierungsverfahrens hinaus lässt sich auch die starke Konvergenz zeigen. Sowohl der Nachweis der starken Konvergenz, als auch der ausführliche Beweis aller bisher gemachten Aussagen ist im Rahmen dieses Seminarvortrages aus Zeitgründen jedoch nicht möglich. Alle Aussagen lassen sich auch für nicht kompakte Operatoren zeigen. Die Beweise der getroffenen Aussagen sind dann jedoch sowohl komplizierter, als auch langwieriger und werden deshalb in diesem Vortrag nicht behandelt.

Literaturverzeichnis

[1] Ingrid Daubechies et al.:
An iterative thresholding algorithm with a sparsity constraint

[2] A.K. Louis: *Inverse und Schlechtgestellte Probleme*