Seminarvortrag aus Angewandter Mathematik
Objektive und subjektive Zusatzinformationen

Michael Kniely

März 2010

1 Allgemeine Problemstellung

Wir betrachten eine Operatorgleichung der Form

$$F(x) = y \quad \text{mit} \quad x \in D \subseteq X, \; y \in Y \tag{1}$$

und das zugehörige inverse Problem; das heißt zu einer gegebenen rechten Seite $y \in Y$ suchen wir ein $x \in D$, sodass die Gleichung in (1) erfüllt ist. Die rechte Seite y nennen wir dabei die Eingangsdaten und die Lösung x den zu identifizierenden Parameter; das Problem in (1) wird dann auch Identifikationsproblem genannt.

In den meisten Fällen ist (1) jedoch nicht korrekt gestellt, das heißt es kann sein, dass es zu bestimmten $y \in Y$ kein $x \in D$ gibt, sodass $F(x) = y$, oder, dass es zwar ein $x \in D$ mit $F(x) = y$ gibt, dieses aber nicht eindeutig bestimmt ist. Schließlich ist es auch noch möglich, dass es zwar zu jedem $y \in Y$ genau ein solch ein gesuchtes x gibt, dieses aber nicht stetig von den Eingangsdaten y abhängt.

Eine Möglichkeit, diese Probleme in den Griff zu bekommen oder wenigstens deren Auswirkungen zu minimieren, besteht in der Verwendung von **Zusatzinformationen** über die Eingangsdaten y und die erwartete Lösung x; Zusatzinformationen über x werden auch **Apriori-Informationen** genannt.

2 Zusatzinformationen

Im Allgemeinen hat man beim Identifikationsproblem in (1) nur eine Näherung y_δ der exakten rechten Seite y zu Verfügung, denn die theoretisch vorhandenen Daten y müssen in irgendeiner Form gemessen werden, wobei meist Messfehler auftreten. Zusatzinformationen über die rechte Seite sind daher Angaben über die Art der Näherung y_δ an die tatsächlichen Werte y; diese werden in **deterministische** und **stochastische Zusatzinformationen** unterteilt, also Informationen, die eine gewisse Schranke für den Fehler $y_\delta - y$ liefern bzw. sich aus wahrscheinlichkeitstheoretischen Überlegungen ergeben.

Ein Beispiel für eine deterministische Zusatzinformation über die Eingangsdaten ist die Beschränkung der Fehlernorm mit einem $\delta > 0$:

$$\|y_\delta - y\|_Y \leq \delta. \tag{2}$$

Betrachtet man die Operatorgleichung $F(x) = y$ mit $F : \mathbb{R}^n \to \mathbb{R}^n$, wobei anstelle des exakten Datenvektors y nur ein gestörter Vektor $y_\varepsilon = y + \varepsilon$ mit $\varepsilon \in \mathbb{R}^n$ bekannt ist, so stellt $\varepsilon \sim \mathcal{N}(0, C)$ eine stochastische Zusatzinformation über die Eingangsdaten dar, wobei $0 \in \mathbb{R}^n$ der Erwartungswertevektor und $C \in \mathbb{R}^{n\times n}$ die Kovarianzmatrix von ε ist. Sind die Fehler in den einzelnen Komponenten unabhängig voneinander, so ist C diagonal; falls zusätzlich noch die einzelnen Messungen die gleiche Varianz σ^2 besitzen, so ist $C = \sigma^2 I$, also Vielfaches der Einheitsmatrix. Die daraus folgende Modellannahme

$$\varepsilon \sim \mathcal{N}(0, \sigma^2 I) \tag{3}$$

wird auch *weißes Rauschen* des Datenfehlers genannt.

Die oben betrachteten Zusatzinformationen schränken den Lösungsbereich des inversen Problems auf jene $x \in D$ ein, deren Bilder $F(x)$ mit den Eingangsdaten verträglich sind. Im Falle der deterministischen Zusatzinformation aus (2) hieße das

$$\|F(x) - y_\delta\|_Y \leq \delta$$

bzw. analog zur stochastischen Zusatzinformation aus (3)

$$\|F(x) - y_\varepsilon\|_2^2 \leq f n \sigma^2,$$

wobei $n\sigma^2 = E(\|\varepsilon\|_2^2)$ der Fehlererwartungswert und $f \in [0.7, 1]$ ein zusätzlicher Faktor ist, der sich empirisch als nützlich erwiesen hat.

3 Apriori-Informationen

Meist gibt es aber auch nach Verwendung deterministischer oder stochastischer Zusatzinformationen noch zu viele $x \in D$, die in dem oben beschriebenen Sinn zu den Daten passen. Daher ist es sinnvoll auch noch Informationen über die zu erwartende Lösung x zu berücksichtigen; dabei unterscheidet man **objektive** und **subjektive Apriori-Informationen**.

Objektive Apriori-Informationen über die Lösung x sind unabhängig vom Bearbeiter des Problems und können z.B. physikalische Tatsachen widerspiegeln, die die Lösung erfüllen muss. Oft ist es möglich diese bereits über den Definitionsbereich D des Operators F in das Problem einfließen zu lassen.

Ist das inverse Problem selbst mit objektiven Apriori-Informationen noch nicht korrekt gestellt, so kann es nützlich sein, auch subjektive Apriori-Informationen - also eigene Vermutungen - heranzuziehen. Diese werden durch ein sogenanntes *Sympathiefunktional* $\Omega : D \to \mathbb{R}$ mit $D \subseteq X$ in die Aufgabe eingebaut. Schließlich werden dann aus allen $x \in D$, die mit den Eingangsdaten verträglich sind, jene ausgewählt, die das Sympathiefunktional Ω minimieren:

$$\Omega(x) \stackrel{!}{=} \text{Min.} \quad \text{mit} \quad x \in D \subseteq X \text{ datenkompatibel.}$$

Daher wird Ω so konstruiert, dass subjektiv passenden x kleine und subjektiv unpassenden x große Werte $\Omega(x)$ zugewiesen werden.

Ein Beispiel für ein solches Funktional ist die möglichst geringe Abweichung von einem Referenzelement $x^* \in X$:

$$\Omega(x) = \|x - x^*\|_X^2.$$

Ist $X = C^n[a,b]$ mit passendem $n \in \mathbb{N}_0$, so verwendet man oft Sympathiefunktionale der Form

$$\Omega(x) = \|x'\|^2_{L^2(a,b)} = \int_a^b (x'(t))^2 dt$$

oder auch Linearkombinationen

$$\Omega(x) = \lambda_0 \|x\|^2_{L^2(a,b)} + \lambda_1 \|x'\|^2_{L^2(a,b)} + \lambda_2 \|x''\|^2_{L^2(a,b)}$$

mit $\lambda_0, \lambda_1, \lambda_2 \geq 0$. Jeweils werden damit weniger stark schwankende Funktionen x bevorzugt. Weiters werden auch Sympathiefunktionale

$$\Omega(x) = \int_a^b x(t) \log \frac{x(t)}{x^*(t)} dt$$

mit einer Referenzfunktion $x^* \in X$ angewandt. Ein Lösungsverfahren mit diesem Funktional nennt man dann *Maximum-Entropie-Regularisierung.*

Bei statistisch beschreibbaren physikalischen Prozessen verwendet man häufig *Bayes'sche Modelle*; bei diesen ist die zu identifizierende Größe x des inversen Problems *randomisierbar*; das heißt, x kann als Realisierung einer Zufallsvariable betrachtet werden. In diesem Fall verwendet man auch die quadratische Form

$$\Omega(x) = (x - \mu)^T B^{-1} (x - \mu)$$

als Sympathiefunktional, sofern die Zufallsvariable vektorwertig ist. Dabei ist μ der Erwartungswertevektor und B die Kovarianzmatrix - beide werden hier als gegeben vorausgesetzt.

Generell sollten subjektive Apriori-Informationen aber sehr zurückhaltend und auch nur bei wirklicher Notwendigkeit eingesetzt werden; denn falsche Vorstellungen und damit unpassende Sympathiefunktionale können sich in der resultierenden Lösung mit beträchtlichen Fehlern niederschlagen.