

Vorwort

Mit der Theorie der partiellen Identifizierung wird Unsicherheit in den beobachteten Daten auf die Schätzung der Modellparameter übertragen. Im Kontext der generalisierten Regression mit Intervalldaten erhält man für die Parameterschätzer so keine skalaren Werte, sondern ebenfalls Intervalle. Im Allgemeinen kann dies als Optimierungsproblem mit Nebenbedingungen formuliert werden. Die Herausforderung besteht hier insbesondere darin, die globalen Extrema zu bestimmen. Für Spezialfälle sind bereits Lösungen aus der Literatur bekannt: Bei der linearen Regression mit skalarer abhängiger Variable kann das Optimierungsproblem analytisch gelöst werden. Für eindimensionale Parameter in generalisierten linearen Modellen wird in der vorliegenden Arbeit ein neuer Ansatz realisiert und untersucht: Optimiert man die Score-Funktion mit festem Parameter über die beobachteten Intervalldaten, so lassen sich dadurch die Extrema des Parameterschätzers bestimmen. Dieser Ansatz liefert zuverlässig die richtige Lösung. Als allgemeine Herangehensweise kann der Parameterschätzer direkt numerisch optimiert werden. Daneben wird ein alternativer Ansatz verwendet: Um das Intervall der zulässigen Schätzer zu erhalten, kann die Score-Funktion als Strafterm in die Zielfunktion integriert werden. Die Methoden werden für das lineare Modell und das generalisierte lineare Modell mit Exponentialverteilung und log-Link umgesetzt. Bei der Untersuchung von Simulationsbeispielen zeigt sich, dass die allgemeinen Verfahren teilweise nur lokale Extrema finden. Da die Zielfunktionen aber mitunter sehr hochdimensional sind, kann die numerische Optimierung nicht an verschiedenen Stellen systematisch neugestartet werden. Auf Grund dieser Problematik werden heuristische Methoden vorgeschlagen, bei denen iterativ in den Ecken der Datenintervalle neugestartet wird. Diese liefern durchwegs die größten Intervalle der zulässigen Parameter und damit die beste Lösung. Die Methode wird auf ein Anwendungsbeispiel aus der Volkswirtschaft angewandt: Vorerst wird ein Regressionsmodell für den Zusammenhang zwischen dem Bruttoinlandsprodukt und den Ausgaben für Forschung und Entwicklung verschiedener Länder entwickelt. Anschließend werden Intervalle für die Daten konstruiert und die entsprechenden Parameterintervalle berechnet. Mit den vorgestellten und praktisch umgesetzten Verfahren lassen sich oft die exakten Lösungen und in allen untersuchten Fällen sehr gute Approximationen finden. Des Weiteren wird demonstriert, dass

die Ansätze grundsätzlich zur Lösung des Problems führen, wenn die Extrema der Optimierungsprobleme richtig bestimmt werden.

Michael Seitz

Intervalldaten und generalisierte lineare Modelle

Seitz, M.

2015, XVII, 110 S. 31 Abb., Softcover

ISBN: 978-3-658-08745-6