Bayesischer Ansatz zur Integration von Expertenwissen in die Immobilienbewertung (Teil 1)

Hamza Alkhatib und Alexandra Weitkamp

Zusammenfassung

Die Immobilienbewertung steht immer wieder im Fokus. Verstärkt werden zuverlässige und überprüfbare Verkehrswerte gefordert. Dies ist derzeit nur in Lagen mit hinreichenden Fällen möglich, in denen das Vergleichswertverfahren in Form einer Regressionsanalyse verwendet wird. Eine Verbesserung des Ergebnisses und eine belastbarere Aussagekraft erhält das Ergebnis durch die Integration von Expertenwissen – dem gutachterlichen Sachverstand. Dieser ist in der Regressionsanalyse auf die Aufstellung des Modells beschränkt.

Ein Bayesischer Ansatz ermöglicht die Integration von Expertenwissen in die Regressionsanalyse. Das Expertenwissen wird als a priori Information modelliert, während die Kauffälle in einer sogenannten Likelihood-Funktion zusammengefasst werden. Über das Bayes-Theorem werden das Expertenwissen und die Daten zusammenführt. Ergebnis ist die a posteriori Information: Regressionskoeffizienten, Konfidenzintervalle u. ä. stehen in verbesserter Form zur Verfügung.

Summary

Currently, subjects dealing with real estate evaluation are in the center of attention. Not least the real estate and finance crisis has shown the importance of correct valuation so that a precise and reliable real estate valuation is increasingly being demand. Normally, in case of sufficient number of purchases, the comparative value method in form of multiple regression analysis will be applied. An improvement of this analysis as well as a meaningful interpretation of the results can be reached by integration of experts' knowledge with the data (in our case the purchases).

The Bayesian approach enables the integration of expert knowledge within the regression analysis. In a first step the expert knowledge is modeled as prior density function, whereas the purchases are combined in another density function – the so called likelihood function. The both density functions are merged by means of Bayes's theorem. The resulting density function is the posterior density, which depend on both: the expert knowledge and the purchases. Based on the later density, different results of interest as estimates of the regression coefficients, confidence regions, and hypothesis testing can be derived.

Schlüsselworte: Wertermittlung, Bayes-Statistik, Expertenwissen, Multiple Lineare Regressionsanalyse, Kauffälle

1 Einleitung

Die präzise und zuverlässige Bereitstellung von Immobilienwerten ist von besonderer sozialer und wirtschaftlicher Bedeutung. Die Immobilien- und Finanzkrise hat gezeigt, welchen Stellenwert die Bewertung einnimmt. Immobilienwerte müssen hohe objektive Bedürfnisse befriedigen. Zudem fordert die deutsche Rechtsprechung eine Genauigkeit für den Marktwert von ± 20 % (BVerfG, Urt. v. 07.11.2006 1 BvL 10/02).

Die Erfüllung dieser Forderung ist einfach, sofern ausreichend verwendbare Daten (passende Kauffälle) vorliegen. Dennoch wird das Expertenwissen in datengetriebene Modelle nicht integriert: Hier wird es »nur« für die Modellierung benötigt. Nicht verwendet wird es für die Modellverbesserung und die Bestimmung der Unsicherheit.

2 Problemdarstellung

Die diesem Beitrag zugrunde liegende Forschungsidee ist es, Expertenwissen – den sogenannten »gutachterlichen Sachverstand« – in datengetriebene Methoden (hier die Regressionsanalyse) einzubinden, um präzisere Ergebnisse zu erhalten und eine gesicherte Aussage zu dem Unsicherheitshaushalt treffen zu können. Daher werden im Folgenden zunächst die theoretischen Grundlagen der Regressionsanalyse vorgestellt. Anschließend wird erläutert, warum der Bayesische Ansatz in der Immobilienbewertung in diesem Beitrag verfolgt wird.

2.1 Die klassische multiple lineare Regression in der Immobilienbewertung

In der Immobilienbewertung werden datengetriebene Modelle verwendet, sofern ausreichend vergleichbare Kauffälle verfügbar sind (Vergleichswertverfahren). Als Analysemethode hat sich die multiple lineare Regression etabliert. Viele Gutachterausschüsse bedienen sich der Regressionsanalyse nach Ziegenbein (1977). Hierbei wird zunächst das Modell aufgestellt: Wahl der Einflussgrößen und Formulierung des funktionalen Zusammenhangs. Ist der funktionale Zusammenhang nicht linear beschreibbar, so werden die Einflussgrößen in einem polynominalen Ansatz formuliert (Ziegenbein 1977, S. 80 ff.). Dieser Ansatz wird hier nicht weiter betrachtet.

Anschließend werden die Anforderung an die Regression überprüft und die Regressionsfunktion berechnet.

Anhand der Residuen wird überprüft, ob alle Anforderung an eine Regressionsanalyse erfüllt sind oder ob iterativ das Modell angepasst werden muss (Ziegenbein 1995).

Das Vergleichswertverfahren, welches auf der multiplen linearen Regressionsanalyse basiert, ist das marktnächste Verfahren in der Immobilienbewertung und ist bei ausreichender Datenlage den anderen Verfahren vorzuziehen. Eine weiterführende Motivation und Diskussion der Regressionsanalyse ist in verschiedensten Standardwerken der Statistik zu finden (z. B. Fahrmeir, Kneib und Lang (2009), Urban und Mayerl (2006)).

Zusammenfassend kann festgehalten werden, dass die linearen Regressionsmodelle die Beziehung der abhängigen Variablen y_i und den unabhängigen Variablen $x_{i2}, \ldots, x_{ik}$ beschreibt. Dies erfolgt in der Form:

$$y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \epsilon_i; \text{ mit } i = 1 \ldots n. \tag{1}$$

Gl. 1 kann in Matrizenschreibweise formuliert werden. Zusammengefasst werden die abhängige Variable (Zielgröße) $\mathbf{y} = [y_1, \ldots, y_n]'$, der Fehlervektor (Residuen) $\boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_n]'$, der $k \times 1$ Parametervektor (Regressionskoeffizienten) $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_k]'$ und die $n \times k$ Matrix der Einflussgrößen $\mathbf{X}$ mit n als Anzahl der (Kauf-) Fälle und k als Anzahl der Parameter (der Parametervektor beinhaltet sowohl die Einflussgrößen als auch die Konstante β_1):

$$\mathbf{X} = \begin{bmatrix} 1 & x_{12} & \ldots & x_{1k} \\ 1 & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \ldots & \vdots \\ 1 & x_{n2} & \ldots & x_{nk} \end{bmatrix}$$

Gl. 1 lässt sich in Matrizenschreibweise als:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2}$$

formulieren.

Die Schätzung der Regressionskoeffizienten in einem linearen Modell erfolgt nach der Methode der kleinsten Quadrate (ordinary least squares / OLS). Dazu bedarf es der Einhaltung allgemeiner Anforderungen: der Erwartungswert der Residuen muss Null ergeben ($E(\boldsymbol{\epsilon}) = \mathbf{0}$). Die Residuen müssen normalverteilt sein ($\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$) und dürfen nicht korrelieren ($Cov(\boldsymbol{\epsilon}) = E(\boldsymbol{\epsilon}'\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$). Die Varianz der Residuen muss konstant für alle Fälle sein (homoskedastische Fehler). Die Zielgröße ist stochastisch und die Matrix der Einflussgrößen muss den vollen Rang ($rg(\mathbf{X})$ = Anzahl der Parameter) besitzen (Fahrmeir, Kneib und Lang (2009), S. 59 ff.).

Der optimale Wert des Vektors $\boldsymbol{\beta}$ ist der Vektor $\hat{\boldsymbol{\beta}}$, der die Summe der Quadrate zwischen dem Zielgrößen-Vektor $\mathbf{y}$ und dem Vektor $\mathbf{X}\boldsymbol{\beta}$ minimiert. Die Lösung der resultierenden Normalgleichung für $\boldsymbol{\beta}$ (vgl. z. B. Koch 1999, Brückner 1976) ergibt:

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \tag{3}$$

$$\text{mit } \mathbf{N} = \mathbf{X}'\mathbf{X} \tag{4}$$

$$\text{und } \mathbf{n} = \mathbf{X}'\mathbf{y} \tag{5}$$

$$\hat{\boldsymbol{\beta}} = \mathbf{N}^{-1}\mathbf{n}. \tag{6}$$

Die Inverse der Normalgleichungsmatrix ergibt die Kofaktormatrix der Parameter:

$$\mathbf{Q}_{\hat{\beta},\hat{\beta}} = \mathbf{N}^{-1} \tag{7}$$

mit der Quadratsumme der Residuen (Brückner 1976, S. 153 ff.):

$$\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{n}. \tag{8}$$

Die Güte der Regression wird nach Fahrmeir, Kneib und Lang (2009) zumeist am Bestimmtheitsmaß B abgelesen:

$$B = 1 - \frac{\sum_{i=1}^{n} \hat{\epsilon}_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}. \tag{9}$$

Hierbei ist $\bar{y}$ das arithmetrische Mittel der Zielgröße (bei n Kauffälle).

Das Bestimmtheitsmaß liegt zwischen 0 und 1. Für $B = 0$ sind keine Korrelationen zwischen Ziel- und Einflussgrößen vorhanden. Für $B = 1$ ist ein strikter linearer funktionaler Zusammenhang festzustellen (Brückner 1976, S. 155 ff.).

Der Test auf Signifikanz der Regression im Falle der multiplen linearen Regressionsanalyse erfolgt über die Analyse der Varianz. Nullhypothese H_0 und Alternativhypothese H_1 sind in diesem Fall:

$$\begin{aligned} H_0: &\quad E(\beta_1) = E(\beta_2) = \ldots = E(\beta_k) = 0 \\ H_1: &\quad E(\beta_i) \neq 0. \end{aligned} \tag{10}$$

Die Teststatistik ist somit zu formulieren als:

$$F_{k,f} = \frac{f}{k} \cdot \frac{B}{1-B} \tag{11}$$

mit n als Anzahl der Fälle, k als Anzahl der Parameter (inkl. Konstante) und mit der Anzahl der Freiheitsgrade als $f = n - k$. Ist die Nullhypothese H_0 zutreffend, so folgt die Teststatistik $F_{p,f}$ der F-Verteilung und die Wahrscheinlichkeitsbeziehung

$$P\left\{\frac{f}{k} \cdot \frac{B}{1-B} > F_{k,f,1-\alpha} | H_0\right\} = \alpha \tag{12}$$

trifft zu (Brückner 1976, S. 153 ff.). α in Gl. 12 gibt die Irrtumswahrscheinlichkeit an, z. B. $\alpha = 5\,\%$.

Zusätzlich werden die einzelnen Regressionskoeffizienten β_i darauf getestet, ob sie sich signifikant von 0 unterscheiden. Dazu werden die Hypothesen wie folgt formuliert:

$$H_0: E(\beta_i) = 0 \;\text{ versus }\; H_1: E(\beta_i) \neq 0. \tag{13}$$

Die Teststatistik basiert auf der Studentverteilung (t-Verteilung). H_0 wird abgelehnt, wenn

$$|t_i| = \frac{|\hat{\beta}_i|}{s\sqrt{q_{ii}}} > t_{f,1-\alpha/2} \tag{14}$$

zutrifft. Dabei sind s der geschätzte Wert der Standardabweichung der Gewichtseinheit und q_{ii} die Werte auf der Diagonalen der Kofaktormatrix der Regressionskoeffizienten $\mathbf{Q}_{\hat{\beta}\hat{\beta}}$ (vgl. Gl. 7).

Die Kauffälle unterliegen einer natürlichen Streuung aufgrund des individuellen Anteils bei den Kaufverhandlungen und der Unterschiede in den Objekten. Diese Streuung kann mit dem Variationskoeffizienten gemessen werden, der das Verhältnis von Standardabweichung der geschätzten Beobachtungen $s_{\hat{y}}$ und Mittelwert $\bar{\hat{y}}$ der Zielgröße angibt:

$$V = \frac{s_{\hat{y}}}{\bar{\hat{y}}}. \tag{15}$$

In Untersuchungen hat sich ergeben, dass (unabhängig von den Kaufpreishöhen) der Variationskoeffizient in den sachlichen Teilmärkten konstant ist; in Niedersachsen beträgt der Variationskoeffizient V = 0,1 für Eigentumswohnungen, V = 0,1 – 0,15 für Ein- und Mehrfamilienhäuser, V = 0,2 für Bauland und V = 0,25 für landwirtschaftliche Flächen (Ziegenbein 2010).

In der Regressionsanalyse müssen nach Ziegenbein (2010) bestimmte Anforderungen eingehalten werden. Es soll eine bestimmte Anzahl an vergleichbaren Fällen (> 15 pro verwendete Einflussgröße) vorliegen. Die Einflussgrößen sollen nicht sehr hoch korreliert sein (vorzugsweise < 0,3 bzw. höchstens < 0,6). In die Regressionsfunktion sollen nur signifikante Einflussgrößen (getestet mit dem t-Test aus Gl. 14) und lediglich Kauffälle des gewöhnlichen Geschäftsverkehres ohne persönliche Verhältnisse eingehen. Auch dürfen keine Einflussgrößen unberücksichtigt bleiben, die einen signifikanten Einfluss auf die Zielgröße haben. Die ideale Regressionsfunktion kann nur durch Einsatz von gutachterlichem Sachverstand hergeleitet werden.

2.2 Hintergrund für einen Bayesischen Ansatz in der Wertermittlung

Dem zweiten Teil dieses Beitrages vorgreifend kann festgehalten werden, dass Sachverständige der Immobilienwertermittlung in der Regel über einen großen Erfahrungsschatz verfügen (Der zweite Teil dieses Beitrages wird sich tiefergehend mit dem Expertenwissen in der Immobilienbewertung auseinandersetzen und insbesondere die unterschiedlichen Sachverständigen vorstellen). Allerdings verbleibt dieser Sachverstand im rein technischen Teil der Bewertung insbesondere der Regressionsanalyse ungenutzt. Er wird zur Aufstellung des Modells verwendet, findet aber in der Regressionsanalyse selbst keinen Eingang. Daher wird hier ein Verfahren vorgestellt, das die Möglichkeit schafft, Expertenwissen in die Regressionsanalyse einzuführen.

Ziel ist es, eine gesichertere Aussage für die Unsicherheit des Verkehrswertes zu erzielen, die durch den unscharfen Markt nicht einfach ist. Der Begriff der »Unsicherheit« wird in diesem Beitrag nicht nach ISO (1995) im Sinne einer Messungenauigkeit verstanden, sondern bezieht sich auf die Genauigkeit der Zielgröße, deren Konfidenzintervalle und mögliche Kovarianzen.

Viele Sachverständige verzichten auf die Angabe einer Genauigkeitsabschätzung für den Verkehrswert (Jester 2006). In der Regressionsanalyse wird die Genauigkeit des Verkehrswertes über die Güte der Regression (Bestimmtheitsmaß B) und die Konfidenzintervalle wiedergegeben.

In diesem Abschnitt wird der Bayesische Ansatz in der multiplen linearen Regressionsanalyse zur Immobilienbewertung kurz vorgestellt. Der Ansatz basiert auf dem Bayes-Theorem, das in Abschnitt 3 eingeführt wird. Weitere Details zur Bayes-Statistik finden sich in vielen Literaturstellen, wie z. B. Koch (2007) und O'Hagan (1994).

In der Bayes-Theorie ist es möglich, a priori Informationen über die Eingangsgrößen zu nutzen. In diesem Fall wird Vorwissen über die Einflussgrößen bzw. Regressionskoeffizienten eingebracht. Dieses a priori Wissen wird mit Hilfe der vollständigen Wahrscheinlichkeitsdichten (PDF) oder kurz Priori-Dichten integriert. Ist der funktionale Zusammenhang der Einflussgrößen linear und kein Vorwissen vorhanden (nicht-informative Priori-Dichte), sind die Ergebnisse des Bayes-Ansatzes und der klassischen Regressionsanalyse identisch.

Ein wichtiges Instrument im Rahmen des Bayes-Ansatzes ist es, mit den Unsicherheiten der abgeleiteten Parameter und deren funktionalem Zusammenhang (z. B. der Regressionsfunktion selbst) umzugehen. Daher schlagen Kacker und Jones (2003) sowie Siebert und Sommer (2004) die Monte-Carlo-Simulation vor, um die Punktschätzung und die Unsicherheiten der Eingangsgrößen in Zusammenhang mit den nicht-linearen multiplen Regressionsfunktionen zu bewerten. Die Analyse der Unsicherheit ist in internationalem Kontext im »Guide to the Expression of Uncertainty in Measurement (GUM)« integriert (ISO 2007). Koch (2008a) bzw. Koch (2008b) entwickelte eine Methode, um den Unsicherheitshaushalt entsprechend GUM in Zusammenhang mit Laserscanning abzuleiten. Eine weitere Studie, die sich mit einem detaillierten Vergleich der Wahrscheinlichkeit und der Fuzzy-Random-Ansätze für den Umgang und die Verbreitung der verschiedenen Unsicherheiten beschäftigt, wird in Alkhatib et al. (2008) eingeführt. Dieser Ansatz basiert auf Kombination von stochastischer und deterministischer Unsicherheit. Der daraus resultierende Bayes-Fuzzy-Ansatz wurde beim terrestrischen Laserscanning angewendet.

Abb. 1 verdeutlicht den Bayes-Ansatz in der Immobilienbewertung. Der Ansatz beruht auf der Integration des Expertenwissens in datengetriebene Modellen, wie der hier vorgestellten multiplen linearen Regression. Das Expertenwissen wird aus Interviews gewonnen. Die Wertermittler schätzen die Veränderung der Verkehrswerte, die durch Variationen des Wertermittlungsobjektes auftreten. Aus diesen Einschätzungen wird ein weiterer Datensatz erstellt, der die eigentlichen Kauffälle nicht enthält. Dieser Datensatz wird als a priori Informationen in einer klassischen multiplen linearen Regressionsanalyse verarbeitet. Als Ergebnis erhält man eine Regressionsfunktion ein-

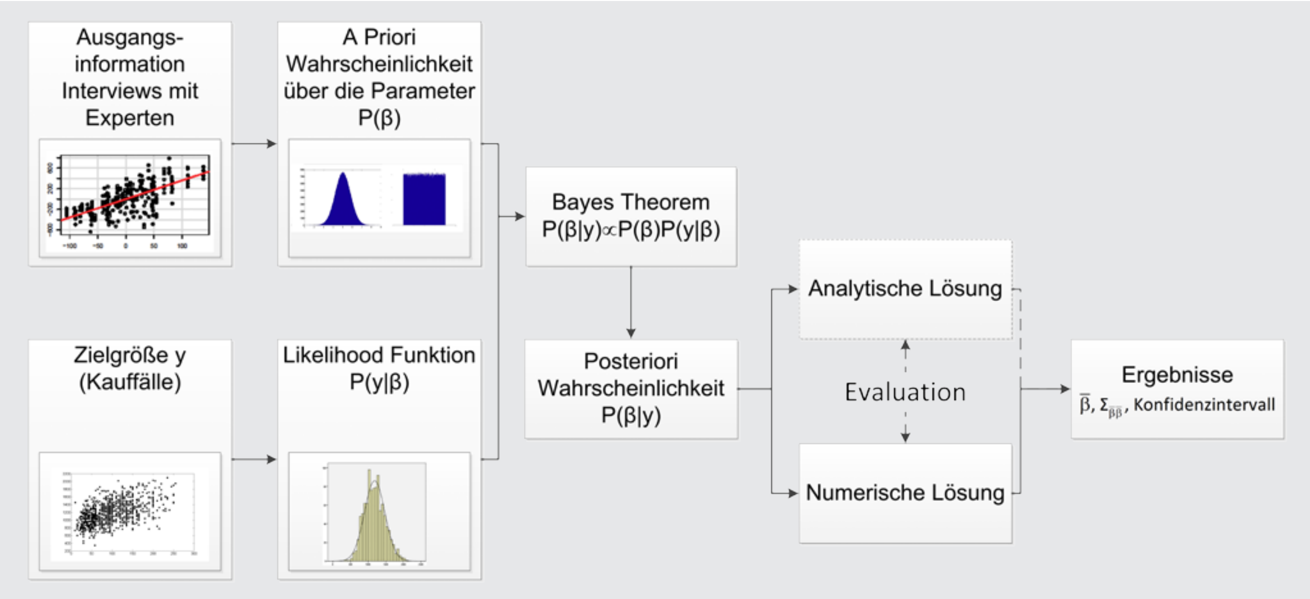

Abb. 1: Der Bayesische Ansatz

schließlich der Aussagen zum Unsicherheitshaushalt. Mittels dieser Regressionsergebnisse werden normalverteilte Priori-Dichten aufgestellt.

Parallel dazu wird aus realen Kauffällen eine Likelihood-Funktion aufgestellt. Die Likelihood-Funktion beinhaltet somit die Informationen, die aus realen Daten gewonnen werden. Die Aufstellung der Likelihood-Funktion kann gleichzeitig mit der Bildung der Priori-Dichten erfolgen: Die Zielgröße »Kaufpreis pro Wohnfläche« (Daten) wird in einer Normalverteilungsfunktion zusammengefasst. Dazu bedient man sich der Erwartungswerte der Zielgröße $(\mathbf{X}\boldsymbol{\beta})$ und des Varianzfaktors der Gewichtseinheit σ^2 multipliziert mit der Gewichtsmatrix der Residuen (in diesem Fall die Einheitsmatrix $\mathbf{I}$): $\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$.

Anschließend werden die Priori-Dichte und die Likelihood-Funktion im Bayes-Theorem zusammengeführt (vgl. Gl. 16). Das Ergebnis ist die sogenannte a posteriori Dichtefunktion. Diese ist aufgrund der Normalverteilungsannahme und der Linearität der multiplen Regressionsfunktion analytisch lösbar. Trifft eine dieser Annahmen nicht zu, so besteht die Möglichkeit, eine Lösung numerisch herbeizuführen. Dazu werden Monte-Carlo-Methoden eingesetzt. Hier wird die numerische Lösung zur Validierung der analytischen Lösung verwendet. Die resultierende Posteriori-Dichte gehört zur Familie der multivariaten Normalverteilung. Hauptergebnis der Posteriori-Dichte ist eine präzisere Punktschätzung der Koeffizienten. Außerdem stehen die Unsicherheiten und eine Schätzung der Vertrauensbereiche zur Verfügung.

Als Ergebnis erhält der Wertermittler aus der a posteriori Dichtefunktion Regressionskoeffizienten, die durch das Expertenwissen verbessert sind. Daneben erlangt er bessere Kenntnisse über den Unsicherheitshaushalt und eine sicherere Abschätzung des Konfidenzintervalls.

Es muss erwähnt werden, dass die angenommene Priori-Dichte einen großen Einfluss auf die resultierende Posteriori-Dichte und damit auf die gesamte Regression hat. Daher kann eine falsche oder schlechte Einschätzung zu einem verzerrten Ergebnis und damit zu Fehlinterpretationen führen. Ist sich der Wertermittler in seiner Einschätzung nicht sicher, besteht allerdings die Möglichkeit, das Gewicht der a priori Information zu verringern oder nicht-informative Dichten zu verwenden (was bedeutet, die Integration der Priori-Dichte zu vermeiden).

Der in Abb. 1 dargestellte Ansatz wird im nächsten Abschnitt theoretisch belegt: Der verwendete Ansatz mit den Teilen Priori-Dichte, Likelihood-Funktion und Posteriori-Dichte wird dargelegt. Ein Leitfaden für die Wertermittlung wird im Abschnitt 4 gegeben. Der Beitrag schließt mit einer methodischen Zusammenfassung und einem Ausblick auf einen weiteren anwendungsorientierten zweiten Teil.

3 Bayesischer Ansatz zur multiplen Regressionsanalyse

In der Bayesischen Interferenz liegt das Interesse darin, etwas über die Regressionskoeffizienten $\boldsymbol{\beta}$ basierend auf den Daten $\mathbf{y}$ zu erfahren. Dazu wird das Bayes-Theorem genutzt:

$$p(\boldsymbol{\beta}|\mathbf{y}) = \frac{p(\boldsymbol{\beta})p(\mathbf{y}|\boldsymbol{\beta})}{p(\mathbf{y})}. \tag{16}$$

Bei dem Term $p(\boldsymbol{\beta}|\mathbf{y})$ handelt es sich um die sogenannte Posteriori-Dichte, den vollständigen Wahrscheinlichkeitsdichten (PDF) der Daten gegeben durch die Parameter des Regressionsmodells. Die Posteriori-Dichte beinhaltet das Wissen über die Parameter im Falle gegebener Daten. $p(\mathbf{y}|\boldsymbol{\beta})$ wird als Likelihood-Funktion bezeichnet und $p(\boldsymbol{\beta})$ als Priori-Dichte.

Der Vektor $p(\mathbf{y})$ als normalisierte Konstante im Nenner der Gl. 16 beinhaltet die Parameter $\boldsymbol{\beta}$ nicht und kann deshalb vernachlässigt werden:

$$p(\boldsymbol{\beta}|\mathbf{y}) \propto p(\boldsymbol{\beta})p(\mathbf{y}|\boldsymbol{\beta}). \tag{17}$$

Die Posteriori-Dichte auf der linken Seite von Gl. 17 verhält sich proportional zur Likelihood-Funktion multipliziert mit der Priori-Dichte. Im nächsten Abschnitt werden die Likelihood-Funktion und die Priori-Dichte im Kontext des multiplen linearen Regressionsmodells mit ergänzender a priori Information und mehreren unabhängigen Variablen entwickelt. In weiteren Ausführungen werden die a priori Größen unterstrichen, z. B. $\underline{s}^2$ für die a priori Varianz, und die a posteriori Größen überstrichen, z. B. $\overline{s}^2$ für die a posteriori Varianz.

3.1 Die Priori-Dichte

Die Priori-Dichte $p(\boldsymbol{\beta})$ enthält Vorwissen bezgl. der Regressionskoeffizienten. Koch (2007) unterscheidet zwischen informativen und nicht-informativen Fällen von Vorwissen. In diesem Modell wird angenommen, dass der Varianzfaktor $h = \frac{1}{\sigma^2}$ des Regressionsmodells eine zufällige unbekannte Variable darstellt.

Die Priori-Dichte $p(\boldsymbol{\beta})$ enthält somit das Wissen über die Parameter $\boldsymbol{\beta}$. Während die informativen Priori-Dichten reales Vorwissen enthalten, fehlt dieses bei den nicht-informativen Priori-Dichten. Die gelungene Anwendung der Bayes-Methode hängt in großem Maße von der geeigneten Wahl des a priori Modells ab.

Die Priori-Dichte, auch genannt die a priori Wahrscheinlichkeit, wird somit aus dem Vorwissen gebildet. Im Fall der Wertermittlung, speziell im Vergleichswertverfahren, kann das Vorwissen über die Eingangsgrößen (hier: Regressionskoeffizienten) z. B. aus einem Experiment bzw. Befragung gewonnen werden. Diese Erkenntnisse über die Regressionskoeffizienten können in die Ermittlung der Regressionsfunktion einschließlich der Parameter $\boldsymbol{\beta}$ und deren Unsicherheitshaushalte einfließen.

3.1.1 Nicht-informative Priori-Dichte

Für nicht-informative a priori Informationen gibt es kein konkretes Vorwissen. Der Bewerter kann oder möchte nicht einschätzen, welche Größe die Parameter einnehmen. Daher wird als Vorwissen für die Regressionskoeffizienten angenommen, dass sie zwischen $-\infty$ und $+\infty$ liegen. Es wird keine Einschränkung durch den Sachverständigen vorgenommen, da es ihm an entsprechenden Informationen mangelt. Die nicht-informative Dichte $p(\boldsymbol{\beta})$ kann dementsprechend nach Koch (2007) wie folgt angenommen werden:

$$p(\boldsymbol{\beta}) \propto \text{const} \;\text{ für }\; -\infty < \beta_i < +\infty \text{ mit } i \in 1,\dots,k \tag{18}$$

und die Priori-Dichte in Form

$$p(h) = \frac{1}{h} \;\text{ für }\; 0 < h < \infty. \tag{19}$$

Diese Vorgehensweise führt zu gleichen Ergebnissen wie sie in der klassischen linearen multiplen Regression erzielt werden, da auch hier kein Vorwissen eingebracht wird. Die Regressionskoeffizienten werden lediglich aus den Daten abgeleitet.

3.1.2 Informative Priori-Dichte

Im Gegensatz dazu enthalten die informativen Priori-Dichten Vorwissen. Es handelt sich um Informationen, die z. B. aus Expertenwissen einer vorgeschalteten Befragung generiert werden. Auch kann das Vorwissen für die informativen Priori-Dichten aus Mittelwerten früherer Daten abgeleitet werden, die in einer angemessenen wissenschaftlichen Art gewonnen wurden.

Sachverständige haben zumeist sehr gute Kenntnisse über die Parameter ihrer Bewertungsverfahren. In der Regressionsanalyse nutzen die Sachverständigen dieses Wissen beispielsweise für die Begrenzung ihrer Stichprobe. Hier definieren sie die Grundgesamtheit, indem sie die Ausreißer aus ihrem gutachterlichen Sachverstand heraus eliminieren. Speziell gibt es Kenntnisse, welchen Wertebereich eine Einflussgröße im gewöhnlichen und unpersönlichen Geschäftsverkehr nicht annehmen kann. Auch werden wertbeeinflussende Größen oft durch den gutachterlichen Sachverstand eingebracht, da datenbasierte Ableitungen vielfach nicht möglich sind. Eine detaillierte Betrachtung des Expertenwissens in der Wertermittlung erfolgt im zweiten Teil des Beitrags.

Im Falle von informativen Priori-Dichten gibt es konkretes Vorwissen, z. B. von Experten; hier empfiehlt Koch (2007, S. 55 ff.), dass die natürliche konjugierte Priori-Dichte als normalverteilt anzusetzen ist – wie in diesem Fall. Eine konjugierte Priori-Dichte hat die Eigenschaft, dass sie nach Multiplikation mit der Likelihood-Funktion eine a posteriori Dichtefunktion erzeugt, die der gleichen Verteilungsfamilie entspringt wie die Priori-Dichte selbst (z. B. könnte aus einer Normalverteilung (a priori) eine Student-Verteilung (a posteriori) resultieren).

Die gemeinsame Priori-Dichtefunktion der Parameter und des Varianzfaktors hat die Form einer Normal-Gamma-Verteilung. Die Normal-Gamma-Verteilung ist zu wählen, wenn natürliche konjungierte Dichten verwendet werden sollen und die Varianz unbekannt ist (vgl. z. B. Koch 2007):

$$\boldsymbol{\beta}, h \sim \mathcal{NG}(\underline{\boldsymbol{\beta}}, \underline{\mathbf{V}}, \underline{b}, \underline{p}). \tag{20}$$

$\underline{\boldsymbol{\beta}}$ enthält die a priori Mittelwerte der Regressionskoeffizienten ($k-$Vektor). $\underline{\mathbf{V}}$ ist die $k \times k$ positiv definite a priori Varianz-Kovarianz-Matrix und entspricht $\mathbf{Q}_{\hat{\beta}\hat{\beta}}$ des klassischen Ansatzes nach Gl. 7. $\underline{p}$ und $\underline{b}$ sind gegeben durch:

$$\underline{p} = (\underline{s}^2)^2/V_{\sigma^2} + 2 \text{ und } \underline{b} = [(\underline{s}^2)^2/V_{\sigma^2} + 1]\underline{s}^2. \tag{21}$$

$\underline{s}^2 = E(\sigma^2)$ ist der Erwartungswert des a priori Varianzfaktors und $V_{\sigma^2} = V(\sigma^2)$ seine Varianz.

Da die Parameter und der Varianzfaktor im Folgenden einzeln zu betrachten sind, werden nun die entsprechenden Verteilungsfunktionen aufgestellt. Ausgehend von der Gl. 20 besitzt die Randverteilung der Parameter – nun ohne den Varianzfaktor h betrachtet – nach Koch (2007) eine Student- oder t-Verteilung der Form:

$$\boldsymbol{\beta} \sim t(\underline{\boldsymbol{\beta}}, \underline{b}\underline{\mathbf{V}}/\underline{p}, 2\underline{p}) \tag{22}$$

Parallel dazu folgt die a priori Randverteilung des Varianzfaktors $\underline{h}$ ebenfalls nach Koch (2007) einer Gamma-Verteilung:

$$\underline{h} = 1/\sigma^2 \sim \mathcal{G}(\underline{b}, \underline{p}). \tag{23}$$

In manchen Fällen bildet das oben eingeführte Priori-Modell das a priori Wissen bezüglich Einflussgrößen und Regressionskoeffizienten nicht vollständig ab. Dies trifft zu, wenn z. B. Restriktionen bzgl. Regressionskoeffizienten integriert werden sollten. Restriktionen einzuführen ist immer dann sinnvoll, wenn die Parameter nach gutachterlichem Sachverstand nicht dem gewöhnlichen und unpersönlichen Geschäftsverkehr entsprungen sind oder Ausreißer darstellen. Diese Größen bilden regelmäßig die extremen Randbereiche der Stichprobe.

Beispielsweise kann ein Gutachter feststellen, dass ein Regressionskoeffizient oder eine Gruppe von Regressionskoeffizienten einem definierten Intervall oder einer Region angehören $\boldsymbol{\beta} \in A$, wobei A die relevante Region oder das relevante Intervall bezeichnet. Dies geschieht häufig durch Beschränkung der Stichprobe in den Einflussgrößen. Werden solche Priori-Dichten in das Bayesische Modell eingeführt, so ist die Auswertung der Posteriori-Dichte im Bayes-Theorem (vgl. Gl. 17) nur anhand von Monte-Carlo-Techniken möglich (Koch 2007, S. 187 ff.). Diese Vorgehensweise wird in diesem Beitrag nicht weiter verfolgt, sondern in nachfolgenden Forschungsarbeiten aufgegriffen.

3.2 Die Likelihood-Funktion

Die Likelihood-Funktion $p(\mathbf{y}|\boldsymbol{\beta})$ ist die Wahrscheinlichkeitsdichte der Beobachtungen, die abhängig von den Parametern des Regressionsmodells ist. Im nächsten Abschnitt wird angenommen, dass die Residuen (der Fehlerterm) der linearen Regressionsfunktion (Gl. 2) einer multiplen Normalverteilung folgen. Das führt zu dem Schluss, dass $p(\mathbf{y}|\boldsymbol{\beta})$ eine normalverteilte Wahrscheinlichkeitsdichte ist. Dies ist mit den Regressionskoeffizienten und deren Varianz verknüpft.

Die Likelihood-Funktion fasst die Informationen aus den realen Daten (Kauffälle) zusammen. Sie dient dazu, die Parameter einer Dichte- bzw. Wahrscheinlichkeitsfunktion zu ermitteln.

In Hussein (2010) wurden die Eingangsgrößen von Renditeobjekten in einem räumlichen Teilmarkt hinsichtlich ihrer Verteilungsfunktion untersucht. Hierbei hat sich ergeben, dass viele Größen linkssteil / rechtsschief ausgeprägt sind, z. B. Bodenwert, Restnutzungsdauer und Reinertrag. Ihre Verteilungsfunktion entsprechen signifikant einer generalisierten Extremwertverteilung, Exponentialverteilung oder Gamma-Verteilung. Lediglich der Liegenschaftszinssatz hat sich als normalverteilt erwiesen. Im Verlauf mehrerer Jahre hat sich die Familie der Verteilungen als konstant erwiesen, wenn auch die Momente wie z. B. der Modus sich veränderten.

In der praktischen Vorgehensweise der Regressionsanalyse werden die schiefsymmetrischen Daten nach Ziegenbein (1977) in eine symmetrische Verteilung transformiert oder die extremen Ränder der Stichprobe als Ausreißer eliminiert, ggf. wird die Einflussgröße als Polynomansatz formuliert.

Darauf wird in diesem Ansatz verzichtet. Stattdessen wird die Likelihood-Funktion für die Zielgröße formuliert. Hier fließen alle Daten ein – in der Wertermittlung sind das die Kauffälle. Zur Einführung der Likelihood-Funktion für das Regressionsmodell nach Gl. 1 müssen einige Annahmen über den Vektor der Residuen $\boldsymbol{\epsilon}$ und die $n \times k$ Matrix der Einflussgrößen $\mathbf{X}$ getroffen werden:

- $\boldsymbol{\epsilon}$ folgt der Normalverteilung mit Mittelwert Null und der Kovarianzmatrix $\sigma^2\mathbf{I}_n$ mit $\sigma^2 = \frac{1}{h}$ als Einheit der Varianz und der Einheitsmatrix $\mathbf{I}$: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, h^{-1}\mathbf{I})$.
- $\mathbf{X}$ hat den vollen Rang und deren Elemente können als deterministisch betrachtet werden.

Mit Hilfe der oben benannten Annahmen und unter Verwendung der Definition einer multivariaten Normalverteilung (siehe z. B. Koch 2007, S. 51 ff.) kann die Likelihood-Funktion wie folgt formuliert werden:

$$p(\mathbf{y}|\boldsymbol{\beta}) = \frac{h^{n/2}}{(2\pi)^{n/2}} \left\{ \exp\left[-\frac{h}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \right\}. \tag{24}$$

Übersichtlicher wird die Likelihood-Funktion nach Gl. 24, wenn sie in der Schreibweise der kleinsten Quadrate formuliert wird:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} \tag{25}$$

und

$$s^2 = \frac{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{f} \quad \text{mit } f = n - k. \tag{26}$$

Werden die Gl. 25 und 26 in Gl. 24 eingesetzt und einige Umstellungen vorgenommen, so kann die Likelihood-Funktion wie folgt formuliert werden:

$$p(\mathbf{y}|\boldsymbol{\beta}, h) = \frac{1}{(2\pi)^{n/2}} \cdot \left\{ h^{1/2} \exp\left[-\frac{h}{2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'\mathbf{X}'\mathbf{X}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right] \right\} \cdot \left\{ h^{f/2} \exp\left[-\frac{hf}{2s^{-2}} \right] \right\}. \tag{27}$$

3.3 Die Posteriori-Dichte

Die Posteriori-Dichte lässt sich aus der Multiplikation der Likelihood-Funktion (Gl. 27) mit der Priori-Dichte (Gl. 18 bzw. Gl. 20) entsprechend des Bayes-Theorems nach Gl. 17 bestimmen. Abhängig von der verfügbaren a priori Information (informativ oder nicht-informativ) werden unterschiedliche Ergebnisse erzielt (vgl. auch Koch 2007, S. 110 ff.).

Die Posteriori-Dichte $p(\boldsymbol{\beta}|\mathbf{y})$ drückt das Wissen über die Regressionskoeffizienten der Regressionsfunktion aus, nachdem das a priori Wissen und die Daten in die Auswertung einbezogen wurden: Mit Hilfe der Daten ist es möglich die Priori-Dichte $p(\boldsymbol{\beta})$ fortzuschreiben. Die resultierende Posteriori-Dichte ist somit eine Kombination aus Daten und datenfreien Informationen (hier: Expertenwissen).

3.3.1 Lösungsweg bei nicht-informativen Priori-Dichte

Im Folgenden wird der Lösungsansatz für die nicht-informative Priori-Dichte aufgezeigt. Das Ergebnis entspricht dem der klassischen Vorgehensweise einer multiplen linearen Regression nach Gl. 3 – 6.

Die nicht-informative Priori-Dichte (vgl. Gl. 18 und 19) ergibt sich mit $\sigma^2 = \frac{1}{h}$ nach:

$$p(\boldsymbol{\beta},h) = p(\boldsymbol{\beta}) \cdot p(h) \propto \text{const} \cdot \frac{1}{h} \propto \frac{1}{h}. \tag{28}$$

Unter Verwendung des Bayes-Theorems (Gl. 17) ergibt sich die Posteriori-Dichte nach:

$$p(\boldsymbol{\beta},h|\mathbf{y}) \propto p(\boldsymbol{\beta},h) \cdot p(\mathbf{y}|\boldsymbol{\beta},h). \tag{29}$$

Werden Gl. 28 und 24 in Gl. 29 eingesetzt, so folgt daraus:

$$p(\boldsymbol{\beta},h|\mathbf{y}) \propto \\ h^{n/2-1} \exp\left(-\frac{1}{2}h(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right). \tag{30}$$

Nach umfangreichen Umformungen der Gl. 30 ergibt sich die Posteriori-Dichte als Normal-Gammaverteilung:

$$\boldsymbol{\beta},h|\mathbf{y} \sim \mathcal{NG}\left(\overline{\boldsymbol{\beta}},\overline{\mathbf{V}},\frac{(n-k)\hat{\sigma}^2}{2},\frac{n-k}{2}\right) \tag{31}$$

mit

$$\overline{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{32}$$
$$\overline{\mathbf{V}} = (\mathbf{X}'\mathbf{X})^{-1} \tag{33}$$

Gl. 31 beschreibt die gemeinsame a posteriori Verteilung der Regressionskoeffizienten und des unbekannten Varianzfaktors, auf die im weiteren Verlauf des Beitrags zurückgegriffen wird. Ergebnis ist eine multivariate Student- bzw. t-Verteilung entsprechend:

$$\boldsymbol{\beta} \sim t(\overline{\boldsymbol{\beta}},\overline{s}^2\overline{\mathbf{V}},n-k). \tag{34}$$

mit $\overline{s}^2$ als a posteriori geschätzter Varianzfaktor (vgl. Gl. 26). Der Bayesische Erwartungswert beträgt: $E(\boldsymbol{\beta}|\mathbf{y}) = \overline{\boldsymbol{\beta}}$.

3.3.2 Lösungsweg bei informativen Priori-Dichte mit normalen konjugierten Priori-Dichten

Wird die informative Priori-Dichte in Gl. 20 und die Likelihood-Funktion aus Gl. 24 im Bayes-Theorem (Gl. 29) zusammengeführt, so ergibt sich:

$$\boldsymbol{\beta},h \sim \mathcal{NG}(\overline{\boldsymbol{\beta}},\overline{\mathbf{V}},\overline{b},\overline{p}) \tag{35}$$

mit

$$\overline{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X}+\underline{\mathbf{V}}^{-1})^{-1}(\mathbf{X}'\mathbf{y}+\underline{\mathbf{V}}^{-1}\underline{\boldsymbol{\beta}}), \tag{36}$$
$$\overline{\mathbf{V}} = (\mathbf{X}'\mathbf{X}+\underline{\mathbf{V}}^{-1})^{-1}, \tag{37}$$
$$\overline{b} = \Big\{2[(\underline{s}^2)^2/V_{\sigma^2}+1]\underline{s}^2(\underline{\boldsymbol{\beta}}-\overline{\boldsymbol{\beta}})' \cdot \\ \underline{\mathbf{V}}^{-1}(\underline{\boldsymbol{\beta}}-\overline{\boldsymbol{\beta}})+(\mathbf{y}-\mathbf{X}\overline{\boldsymbol{\beta}})' \cdot \\ (\mathbf{y}-\mathbf{X}\overline{\boldsymbol{\beta}})\Big\}/2, \tag{38}$$
$$\overline{p} = (n+2((\underline{s}^2)^2)(V_{\sigma^2}+4)/2. \tag{39}$$

Der vorherige Ausdruck beschreibt die gemeinsame a posteriori Verteilung. Eine umfassende Ableitung der Verteilung, gegeben in Gl. 35, und deren Parameter, gegeben in den Gl. 36 – 39, erfolgt in Koch (2007, S. 118 f.). Der Varianzfaktor h kann integriert werden. Als Ergebnis ergibt sich eine multivariate t-Verteilung:

$$\boldsymbol{\beta} \sim t\left(\overline{\boldsymbol{\beta}},\frac{\overline{b}\cdot\overline{\mathbf{V}}}{\overline{p}},2\overline{p}\right). \tag{40}$$

Daraus folgt aus der t-Verteilung

$$E(\boldsymbol{\beta}|\mathbf{y}) = \overline{\boldsymbol{\beta}} \tag{41}$$
$$Var(\boldsymbol{\beta}|\mathbf{y}) = \overline{s}^2\overline{\mathbf{V}} \tag{42}$$

mit

$$\overline{s}^2 = \left(n+2\frac{(\underline{s}^2)^2}{V_{\sigma^2}}+2\right)^{-1} \cdot \\ \Big(2((\underline{s}^2)^2/V_{\sigma^2}+1)\underline{s}^2+(\underline{\boldsymbol{\beta}}-\overline{\boldsymbol{\beta}})\underline{\mathbf{V}}^{-1}(\underline{\boldsymbol{\beta}}-\overline{\boldsymbol{\beta}})+ \\ (\mathbf{y}-\mathbf{X}\overline{\boldsymbol{\beta}}'(\mathbf{y}-\mathbf{X}\overline{\boldsymbol{\beta}})\Big). \tag{43}$$

Die Wahrscheinlichkeiten aus der inverse Gammaverteilungen bedeuten:

$$h|\mathbf{y} = \mathcal{IG}(\overline{b},\overline{p}) \tag{44}$$

Der Erwartungswert von $h = \frac{1}{\sigma^2}$ ergibt sich zu:

$$E(h|\mathbf{y}) = \frac{\overline{b}}{\overline{p}-1} \text{ und} \\ Var(h|\mathbf{y}) = \frac{2(\overline{s}^2)^2}{n+2(\underline{s}^2)^2/V_{\sigma^2}}. \tag{45}$$

3.4 Bayes-Konfidenzregion

Es ist sinnvoll, nicht nur eine Punktschätzung der Parameter (Mittelwert und Standardabweichung) anzugeben, sondern auch einen Vertrauensbereich für die geschätzten Größen zu ermitteln. Wenn C eine Region innerhalb von Ω darstellt und unter Annahmen, dass dem Datensatz $C \subseteq \Omega$ einer Irrtumswahrscheinlichkeit von α zugrunde liegt, gilt unter Berücksichtigung der Posteriori-Dichte $p(\beta_i|\mathbf{y})$:

$$p(\beta_i \in C|\mathbf{y}) = \int_C p(\beta_i|\mathbf{y})d\beta_i = 1 - \alpha. \tag{46}$$

Die Hauptaufgabe besteht darin, eine Teilmenge zu finden, in der die Parameter mit einer hohen Wahrscheinlichkeit enthalten sind, z. B. 95 %. Unter der Annahme von beispielsweise $p(\beta_i|\mathbf{y}) \sim \mathcal{N}(0,1)$ ergibt sich für eine Sicherheitswahrscheinlichkeit von 95 % ein Annahmebereich von $[-1.96, 1.96]$. Diese Wahl des Vertrauensbereiches definiert die kleinste Fläche bzw. das kürzeste zuverlässige Intervall. Eine solche Wahl wird als *Highest Posterior Density Interval* (kurz HPDI) bezeichnet. Detailliertere Ausführungen finden sich in Koch (2007, S. 71). In dieser Anwendung als konjungiertes lineares Regressionsmodell kann das HPDI für die β_i berechnet werden, indem die Wahrscheinlichkeiten der $t-$Verteilung nach Gl. 34 für nicht-informative und Gl. 40 für informative Priori-Dichten verwendet werden:

$$\begin{aligned} &p(b_{low} \leq \beta_i \leq b_{high}) = \\ &(\overline{\beta_i} - K \leq \beta_i \leq \overline{\beta_i} + K) = 1 - \alpha \end{aligned} \tag{47}$$

mit $K = t_{inv}(1 - \alpha/2, f) \cdot \sqrt{diag(\overline{\mathbf{V}})}$; t_{inv} stellt die Inverse der kumulativen Student'schen Dichtefunktion mit f Freiheitsgraden und einer Sicherheitswahrscheinlichkeit von $(1 - \alpha/2)$ dar.

3.5 Bayesische Prädiktion

Gutachter sind in der Regel daran interessiert, Objekte mit Hilfe der Regressionsfunktion zu bewerten und den Verkehrswert als abhängige Variable zu ermitteln: $\mathbf{y}^*$. Diese Prädiktion soll auf der sogenannten prädiktiven Dichte $p(\mathbf{y}|\mathbf{y}^*)$ basieren. Diese Dichte lässt sich aus den gemeinsamen Dichten ableiten (Koch 2007):

$$p(\mathbf{y}|\mathbf{y}^*) = \int p(\mathbf{y}^*|\mathbf{y}, \boldsymbol{\beta})p(\boldsymbol{\beta}|\mathbf{y})d\boldsymbol{\beta}. \tag{48}$$

Die Schätzung der Zielgröße $\mathbf{y}^* = (y_1^*, \ldots, y_T^*)'$ für T unbekannte Werte (hier: Prädiktion von Verkehrswerten) ergibt sich nach:

$$\mathbf{y}^* = \mathbf{X}^*\bar{\boldsymbol{\beta}} + \boldsymbol{\epsilon}^*. \tag{49}$$

Die Residuen $\boldsymbol{\epsilon}^*$ entspringen derselben Verteilung und denselben Momenten wie $\boldsymbol{\epsilon}$ und die Designmatrix $\mathbf{X}^*$ hat die Größe $T \times k$.

$$\mathbf{y}^*|\mathbf{y} \sim t\left(\mathbf{X}^*\overline{\boldsymbol{\beta}}, \frac{\overline{b}}{\overline{p}}\left(\mathbf{I}_T + \mathbf{X}^{*\prime}\overline{\mathbf{V}}\mathbf{X}^*\right), 2\overline{p}\right). \tag{50}$$

Diese Gleichung wird verwendet, um die Zielgröße in einem linearen Regressionsmodell mit konjungierter Priori-Dichte zu schätzen. Werden nicht-informative Priori-Dichten verwendet, so ergeben sich die geschätzten Größen nach:

$$\mathbf{y}^*|\mathbf{y} \sim t\left(\mathbf{X}^*\boldsymbol{\beta}, \overline{s}^2\overline{\mathbf{V}}, n-k\right). \tag{51}$$

3.6 Numerische Monte-Carlo-Methode

Die Schätzung der Parameter, die Prädiktion neuer unbekannter Zielgrößen (abhängiger Variablen) und die Schätzung der Konfidenzintervalle wurden in den vorherigen Abschnitten vorgestellt. Basierend auf den Posteriori-Dichten wurden die Schätzungen analytisch ermittelt.

Soll eine nicht-lineare Funktion mit Regressionskoeffizienten $g(\boldsymbol{\beta})$ abgeleitet werden, statt $\boldsymbol{\beta}$ selbst zu schätzen, bedarf es Monte-Carlo-Techniken, um den Mittelwert oder höhere Momente abzuleiten (vgl. u. a. Alkhatib 2007, Alkhatib et al. 2008, Koch 2008a, Beer 2006):

$$\hat{E}\left(g(\boldsymbol{\beta}|\mathbf{y})\right) = \hat{g}(\boldsymbol{\beta}) = \frac{1}{M}\sum_{i=1}^{M} g(\boldsymbol{\beta}^{(i)}), \tag{52}$$

oder

$$\begin{aligned} &Var(\boldsymbol{\beta}|\mathbf{y}) = \\ &\frac{1}{M}\sum_{i=1}^{M}\left(g(\boldsymbol{\beta}^{(i)}) - \hat{g}(\boldsymbol{\beta})\right)\left(g(\boldsymbol{\beta}^{(i)}) - \hat{g}(\boldsymbol{\beta})\right)'. \end{aligned} \tag{53}$$

Werden zufällige Stichproben aus der Posteriori-Dichte für $\boldsymbol{\beta}$ generiert, so werden diese aus der $t-$Verteilung nach Gl. 40 für die informativen und nach Gl. 34 für die nicht-informativen Fälle gewonnen. Dargestellt ist lediglich der Fall für informative Dichten (vgl. Abschnitt 3.5):

Schritt 1: Generiere zufällige Stichproben $\boldsymbol{\beta}^{(i)}$ aus der Posteriori-Dichte, Gl. 40, unter Verwendung eines Zufallsgenerators für die multivariate t-Verteilung.

Schritt 2: Evaluiere die nicht-lineare Funktion $g(\boldsymbol{\beta}^{(i)})$ und speichere die Ergebnisse.

Schritt 3: Wiederhole Schritte 1 und 2 M-mal.

Schritt 4: Verwende die Stichproben, um den Mittelwert der jeweiligen Regressionskoeffizienten unter Verwendung von Gl. 52 und den Mittelwert der dazugehörigen Varianzen unter Verwendung von Gl. 53 zu berechnen.

4 Leitfaden für die Wertermittlung

Abb. 2 versteht sich als Leitfaden für den Einsatz der Bayesische Methoden in der Wertermittlung. Ausgehend von den Daten ergeben sich die Möglichkeiten, äquivalent zur klassischen Regression, ohne Vorwissen nicht-informative Prior-Dichten einzuführen oder das Expertenwissen über informative Priori-Dichten zu integrieren.

In der analytischen Lösung ergeben sich dementsprechend Posteriori-Dichten für die nicht-informativen wie auch für die informativen Priori-Dichten. Ergebnisse sind u. a. die Regressionskoeffizienten und die Kofaktormatrix. Lässt sich das Bayes-Theorem analytisch nicht lösen oder soll die analytische Lösung evaluiert werden, so kann mit Hilfe von Monte-Carlo-Methoden eine numerische Lösung herbeigeführt werden. Die empirische Schätzung der geschätzten Parameter nach Monte-Carlo-Methoden in Abb. 2 ergibt sich durch das Einsetzen der Funktion $g(\boldsymbol{\beta}) = \boldsymbol{\beta}$ (vgl. Abschn. 3.6).

Als Bayesische Prädiktion können in einem letzten Schritt für die nicht-informativen wie auch für die informativen Priori-Dichten Schätzwerte berechnet werden. Die Ergebnisse für die nicht-informativen Priori-Dichten entsprechen denen der klassischen Regressionsanalyse.

5 Methodisches Fazit und Ausblick

Die Wertermittlung unterliegt besonderen Anforderungen an Zuverlässigkeit und Genauigkeit, die mit dem Bayesischen Ansatz verbessert werden können. Vorteil des Ansatzes ist die Möglichkeit, Expertenwissen und Daten – somit die Kauffälle – kombiniert auszuwerten. Damit wird die Beschränkung aufgehoben, das Expertenwissen nur für die Aufstellung des Modells verwenden zu können. Der Ansatz lässt es zu, sowohl informative Priori-Dichten (Vorwissen) als auch nicht-informative Priori-Dichten (ohne konkretes Vorwissen) einzubringen. Somit ist der Sachverständige nicht zwingend verpflichtet, Vorwissen in die Auswertung einfließen zu lassen.

Nicht zu vernachlässigen ist die Auswahl der Priori-Dichten: Eine nicht zutreffende bzw. falsche Wahl führt

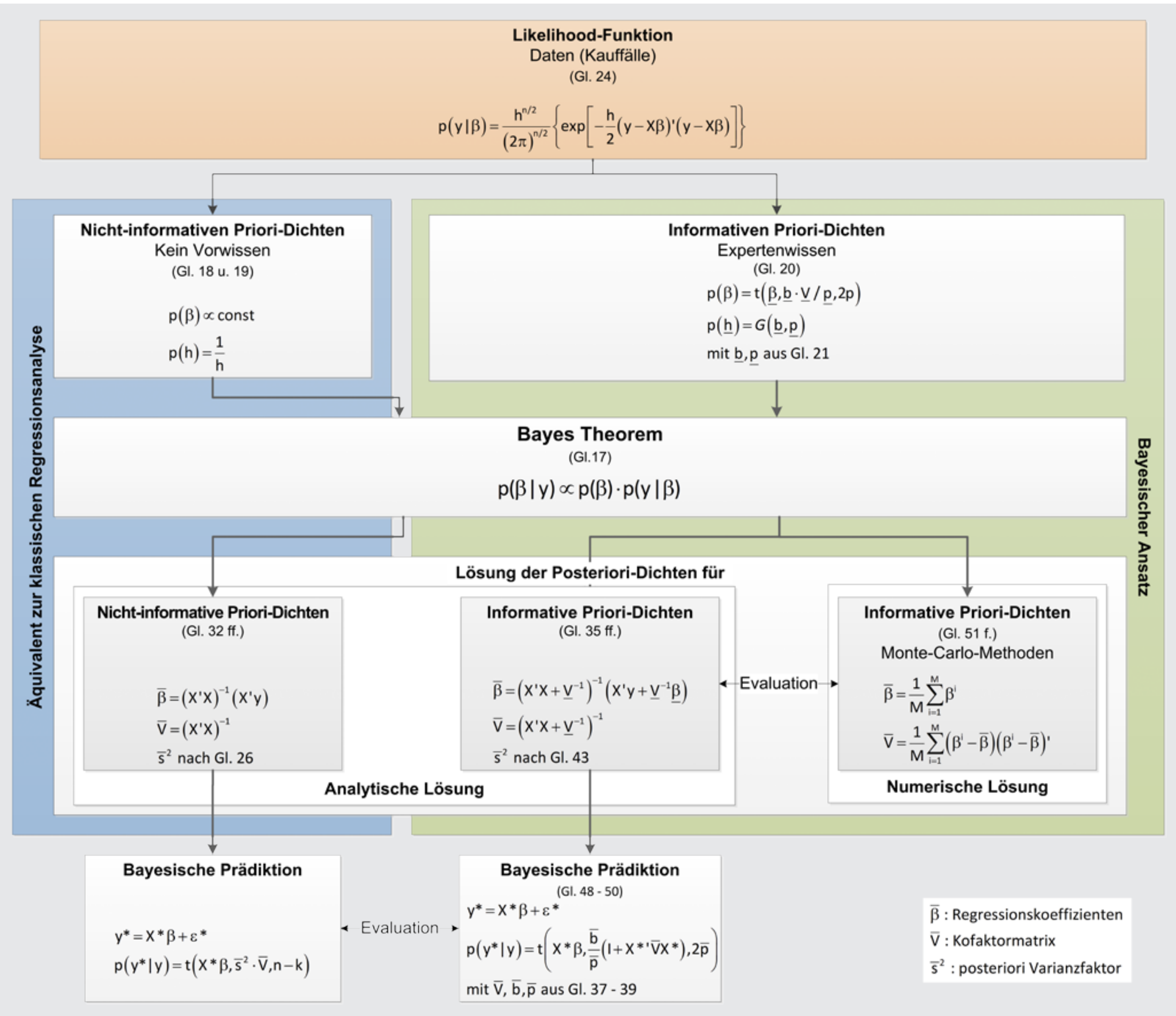

Abb. 2: Leitfaden für die Wertermittlung

zu Verfälschungen der Ergebnisse und damit verbunden zu Fehlinterpretationen. Hier ist der Sachverständige gehalten, eine zutreffende Wahl vorzunehmen oder ggf. auf nicht-informative Priori-Dichten auszuweichen.

Informationsverluste, die durch die klassische Transformation der Ziel- oder Einflussgrößen ggf. entstehen können, existieren im Bayesischen Ansatz nicht. Selbst nicht-lineare Zustände, welche im klassischen Regressionsmodell durch Approximation nicht optimal abgebildet werden können, werden im Bayesischen Ansatz mit wenig Aufwand gelöst.

Weitere Forschungsarbeiten sollen sich mit der Vorverarbeitung von Daten beschäftigen. Diese bedarf vieler manueller Schritte wie beispielsweise der Definition von Beschränkungen der Stichprobe, so dass Ausreißer eliminiert werden. Zukünftig sollen diese Restriktion in das Bayesische Modell eingebunden werden. Dies erfolgt in zusätzlichen Gleichungen oder ggf. Ungleichungen (wie z. B. Verkehrswert ≥ 0). Dies führt allerdings zu einer analytisch nicht lösbaren Posteriori-Dichte. Hier werden die Monte-Carlo-Verfahren ihren Einsatz als effizientes Werkzeug zur Berechnung der Werte finden.

Nach der Literatur ist zu erwarten, dass die klassische Regressionsanalyse mit Beschränkungen dem Bayesischen Ansatz unterliegen wird, da im klassischen Modell dies nur mit Hilfe von Vereinfachungen realisierbar ist. In weiteren Studien soll zudem eine volle Gewichtsmatrix zur Integration der Korrelationen in das Bayesische Modell aufgenommen werden. Die Korrelationen wurden in diesem Beitrag vernachlässigt. Auch diese können aufgrund von Expertenwissen Eingang in das Modell finden.

Im zweiten Teil dieses Beitrags (siehe S. 103 ff.) wird die oben beschriebene Vorgehensweise in dem räumlichen Teilmarkt Osnabrück für Ein- und Mehrfamilienhäuser angewendet. Durch Befragung von Mitgliedern des Gutachterausschusses Osnabrück ist es möglich, Expertenwissen abzuleiten. Mit Hilfe des oben beschriebenen Bayesischen Ansatzes wird das Expertenwissen mit den vorhandenen Kauffällen kombiniert. Dadurch wird eine Verbesserung des Unsicherheitshaushaltes der Regressionskoeffizienten erreicht.

Literatur

Alkhatib, H.: On Monte Carlo Methods with Applications to the Current Satellite Missions. Dissertation, Schriftenreihe des Instituts für Geodäsie und Geoinformation, Landwirtschaftliche Fakultät, Universität Bonn, Nr. 5, 2007.

Alkhatib, H. et al.: Comparison of Sequential Monte Carlo Filtering with Kalman Filtering for Nonlinear State Estimation. In: Ingensand, H. und Stempfhuber (Hrsg.): Proceedings of the 1st International Conference on Machine Control and Guidance. Zurich, Switzerland, 2008, S. 132–142.

Beer, M.: Sampling without probabilistic model. In: Muhanna, R. L. und Mullen, R. L. (Hrsg.): Proceedings of the NSF Workshop on Reliable Engineering Computing – Modeling Errors and Uncertainty in Engineering Computations. 2006.

Brückner, R.: Mathematische Statistik bei der Ermittlung von Grundstückswerten – Lehrbriefe und Vorlesungen zum Kontaktstudium des Geodätischen Instituts 1976. In: Wissenschaftliche Arbeiten der Lehrstühle Geodäsie, Photogrammetrie und Kartographie an der Technischen Hochschule Hannover, Nr. 65 Niedersächsisches Landesververwaltungsamt – Landesvermessung, 1976.

Fahrmeir, L., Kneib, T. und Lang, S.: Regression: Modelle, Methoden und Anwendungen. 2. Auflage. Heidelberg: Springer, 2009, Statistik und ihre Anwendungen.

Hussein, Z.: Geodätische Auswertemethoden zur Ertragswertermittlung. Diplomarbeit am Geodätischen Institut der Leibniz Universität Hannover, unveröffentlicht, September 2010.

ISO: Guide to the expression of uncertainty in measurements (GUM). Geneva: International Organization for Standardization, 1995.

ISO: Propagation of distributions using a Monte Carlo method. In: Evaluation of measurement data – Supplement 1 to the »Guide to the expression of uncertainty in measurement«. Geneva: Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, 2007.

Jester, S.; Roesch, G.: Der Verkehrswert als Näherungswert. Grundstücksmarkt und Grundstückswert, 03 2006, S. 157–161.

Kacker, R. und Jones, A.: On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent. Metrologia, 40 2003, S. 235–248.

Koch, K.-R.: Parameter Estimation and Hypothesis Testing in Linear Models. Berlin – Heidelberg – New York: Springer, 1999, 2nd Ed.

Koch, K.-R.: Introduction to Bayesian Statistics. 2. Auflage. Berlin – Heidelberg – New York: Springer, 2007.

Koch, K.-R.: Evaluation of uncertainties in measurements by Monte Carlo simulations with an application for laserscanning. J Applied Geodesy, 2 2008a, S. 67–77.

Koch, K.-R.: Determining uncertainties of correlated measurements by Monte Carlo simulations applied to laserscanning. J Applied Geodesy, 2 2008b, S. 139–147.

O'Hagan, A.: Bayesian Inference, Kendall's Advanced Theory of Statistics, Vol. 2B. Wiley, New York, 1994.

Siebert, B. R. L. und Sommer, K.-D.: Weiterentwicklung des GUM und Monte-Carlo-Techniken. Technisches Messen, 71 2004, S. 67–80.

Urban, D. und Mayerl, J.: Regressionsanalyse: Theorie, Technik und Anwendung. 2. Auflage. Wiesbaden: VS Verl. für Sozialwissenschaften, 2006, Studienskripten zur Soziologie, ISBN 3531337394.

Ziegenbein, W.: Zur Anwendung multivariater Verfahren der mathematischen Statistik in der Grundstückswertermittlung. Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 77, 1977.

Ziegenbein, W.: Programmgesteuerte Regressionsanalyse und Vergleichswertermittlung im Programmsystem AKS. Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltungen, 4 1995, S. 243–248.

Ziegenbein, W.: Immobilienwertermittlung. In: Kummer, K. und Frankenberger, J. (Hrsg.): Das deutsche Vermessungs- und Geoinformationswesen. Wichmann-Verlag, 2010, S. 421–468.

Anschrift der Autoren

Dr.-Ing. Hamza Alkhatib
Geodätisches Institut der Leibniz Universität Hannover
Auswertemethoden
Nienburger Straße 1, 30167 Hannover
Tel. 0511 762-2464, Fax 0511 762-2468
alkhatib@gih.uni-hannover.de

Dr.-Ing. Alexandra Weitkamp
Geodätisches Institut der Leibniz Universität Hannover
Flächen- und Immobilienmanagement
Nienburger Straße 1, 30167 Hannover
Tel. 0511 762-2406, Fax 0511 762-2468
weitkamp@gih.uni-hannover.de