Statistical Modelling Based on Structured Additive Regression

Habilitationsvortrag von

Thomas Kneib

Institut für Statistik und Ökonometrie
Georg-August-Universität Göttingen

Institut für Statistik
Ludwig-Maximilians-Universität München

16.1.2009

Übersicht

- Anmeldung im November 2006
- Einreichung der Habilitationsschrift am 30.6.2008
- Fachmentorat:
  - Ludwig Fahrmeir (Ludwig-Maximilians-Universität München)
  - Thomas Augustin (Ludwig-Maximilians-Universität München)
  - Göran Kauermann (Universität Bielefeld)
- Gutachter:
  - Ludwig Fahrmeir (Ludwig-Maximilians-Universität München)
  - Göran Kauermann (Universität Bielefeld)
  - David Ruppert (Cornell University)

Übersicht

- Strukturiert Additive Regression
- Bayesianische Regularisierung
- Inferenz I: Gemischte Modelle
- Inferenz II: Markov Chain Monte Carlo Verfahren
- Inferenz III: Boosting
- Zusammenfassung und Diskussion

Strukturiert Additive Regression

- Ziel: Flexible Klasse von Regressionsmodellen, die unter Anderem nichtparametrische und räumliche Effekte beinhalten.

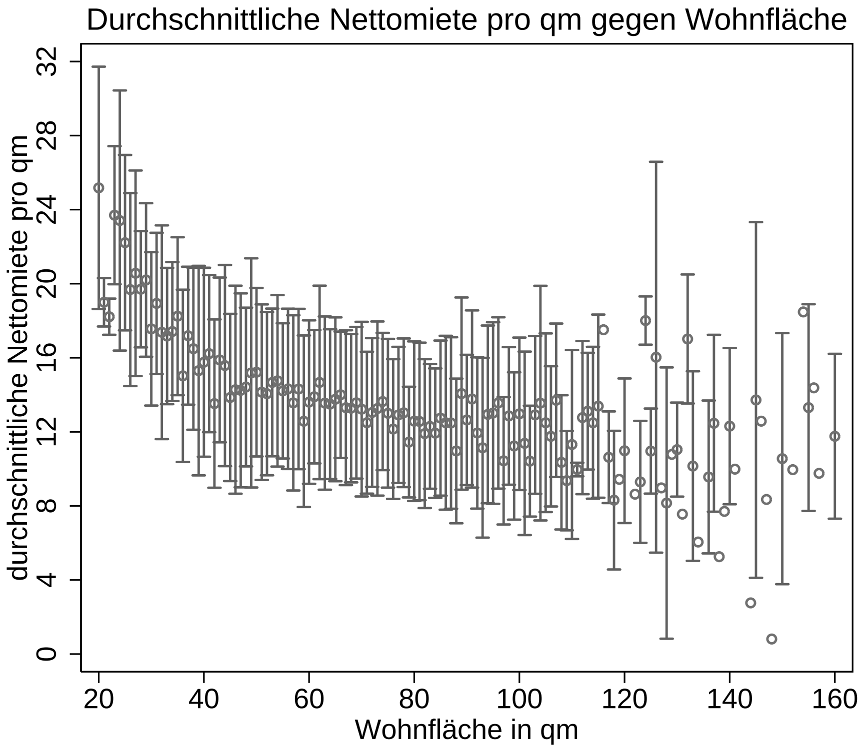
- Illustration am Beispiel des Münchner Mietspiegels.

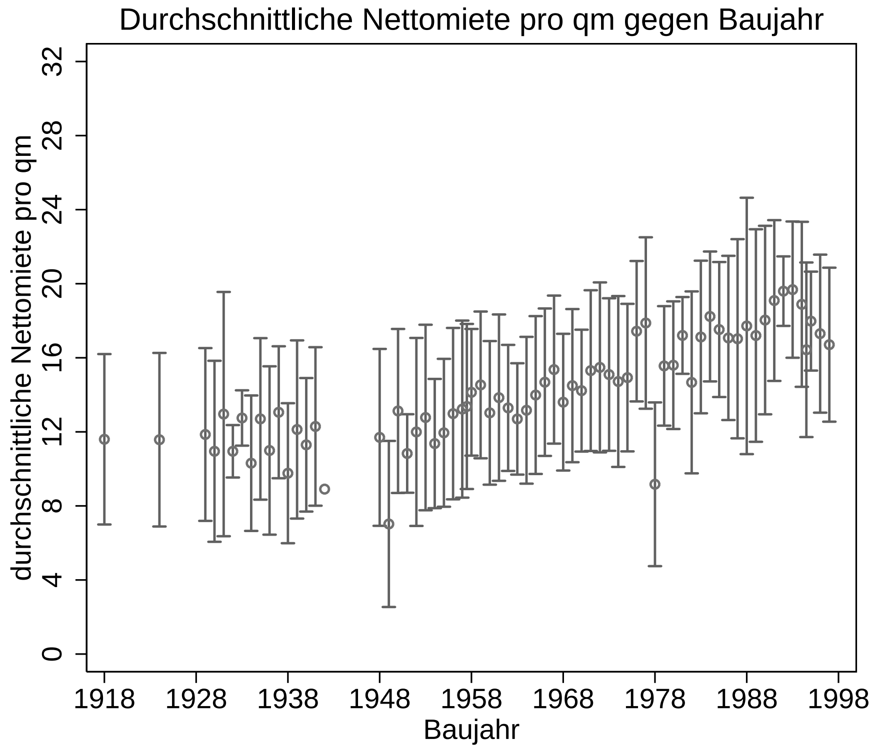
- Erstellt zur Bestimmung der ortsüblichen Vergleichsmiete.

- Verwendung von Regressionsmodellen mit Nettomiete pro Quadratmeter als Zielvariable.

- Übliche parametrische Regressionsmodelle sind aus den folgenden Gründen ungeeignet:

- Nichtlineare Effekte beispielsweise der Wohnfläche oder des Baujahrs.





- Räumliche Information zusätzlich zur Kovariablen "Experteneinschätzung der Wohnlage".

Bezirksviertel in München

- Große Liste kategorialer Kovariablen (ca. 250).

- Einbezug aller Kovariablen führt zu Kollinearitäten und großen Standardfehlern der Schätzer.

  $\Rightarrow$ Regularisierung notwendig (als Alternative zu Variablenselektionsverfahren).

- Mögliches Regressionsmodell für die Nettomiete pro qm nm:

$$nm = f_1(wohnflaeche) + f_2(baujahr) + f_3(bezirksviertel) + x'\beta + \varepsilon.$$

  wobei $x'\beta$ die Effekte eines potenziell hochdimensionalen Vektors von Kovariablen beinhaltet.

- Allgemeine Modellstruktur für strukturiert additive Regressionsmodelle:

$$\eta = f_1(z) + \ldots + f_p(z).$$

- $f_1(z), \ldots, f_p(z)$ sind Funktionen verschiedener Form und z bezeichnet einen generischen Vektor aller Kovariablen, z.B.
  - Lineare Effekte mit oder ohne Regularisierung: $f_j(z) = x\beta$.
  - Nichtlineare, glatte Effekte metrischer Kovariablen: $f_j(z) = f(x)$
  - Interaktionsoberflächen $f_j(z) = f(x_1, x_2)$.
  - Räumliche Effekte: $f_j(z) = f_{\text{spat}}(s)$.
  - Variierende Koeffizienten: $f_j(z) = x_1 f(x_2)$.
  - Räumlich variierende Effekte: $f_j(z) = x f_{\text{spat}}(s)$ oder $x_1 f(x_2, x_3)$.
  - Random Intercepts: $f_j(z) = b_c$ mit Cluster-Index c.
  - Random Slopes: $f_j(z) = x b_c$ mit Cluster-Index c.

Bayesianische Regularisierung

- Allgemeines Ziel der Regularisierung von Schätzproblemen: Schätzung “gutartiger” machen, als sie eigentlich ist.

- Beispiele:
  - Penalisierte Schätzung von Spline-Koeffizienten: Regularisierung durch Norm im Funktionenraum.
  - Räumliche Glättung: Regularisierung durch Annahmen zum Einfluss zwischen benachbarten Regionen.
  - Hochdimensionale Kovariablen: Regularisierung durch geeignete Normen der Regressionskoeffizienten um sparsame Modelle zu erzielen.

- Verschiedene Ziele: Glattheitseigenschaften bzw. sparsame Modellstruktur.

- Frequentistischer Regularisierungsansatz: Ergänze einen Strafterm zur Likelihood bzw. allgemeiner dem zur Schätzung verwendeten Kriterium.

- Beispiel: Lineares Modell

$$y = X\beta + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I).$$

- Für hochdimensionale Kovariablenvektoren wird die Kleinste Quadrate (KQ)-Schätzung von β instabil oder sogar unmöglich.

  $\Rightarrow$ Ergänze das KQ-Kriterium um einen quadratischen Strafterm (Ridge-Regression)

$$KQ_{\text{pen}}(\beta) = (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^{p} \beta_j^2 \rightarrow \min_{\beta}.$$

- Lösung in geschlossener Form: Penalisierter Kleinste Quadrate (PKQ)-Schätzer

$$\hat{\beta} = (X'X + \lambda I)^{-1} X'y.$$

- Bayesianische Formulierung des linearen Modells:

$$y = X\beta + \varepsilon, \qquad \beta \sim \mathrm{N}(0, \tau^2 I).$$

- Man erhält die Posteriori-Verteilung

$$p(\beta|y) \propto \exp\left(-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right) \exp\left(-\frac{1}{2\tau^2}\beta'\beta\right)$$

- Maximierung der Posteriori-Verteilung ist äquivalent zur Minimierung des penalisierten KQ-Kriteriums

$$(y - X\beta)'(y - X\beta) + \lambda\beta'\beta$$

wobei der Glättungsparameter gegeben ist durch das Signal-Rauschen-Verhältnis

$$\lambda = \frac{\sigma^2}{\tau^2}.$$

- Der Posteriori Modus ist äquivalent zum penalisierten KQ-Schätzer.

- Analog für allgemeinere Formen von Priori-Verteilungen:

  – Penalisierte Log-Likelihood:

$$l_{\text{pen}}(\beta) = l(\beta) - \text{pen}(\beta).$$

  – Posteriori:

$$p(\beta|y) = p(y|\beta)p(\beta).$$

- Insgesamt erhält man die Beziehung

Penalisierung $\equiv$ logarithmierte Priori-Verteilung.

- Bayesianische Regularisierung entspricht der Annahme einer informativen Priori-Verteilung.

- Wesentliche Klasse in strukturiert additiven Regressionsmodellen: Quadratische Strafterme

$$\text{pen}(\beta) = \lambda \beta' K \beta$$

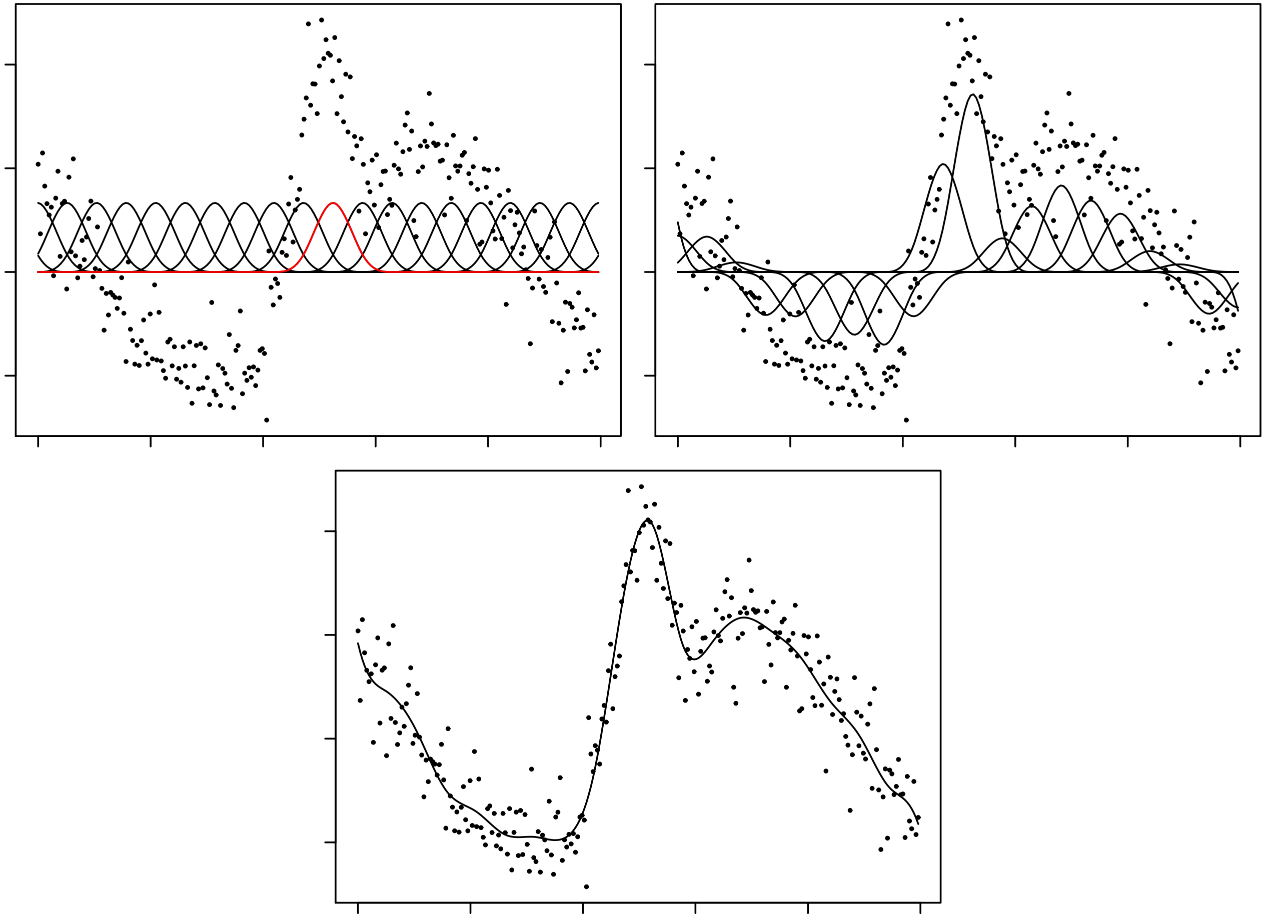
bzw. multivariate Gauß-Prioris

$$p(\beta) \propto \exp\left(-\frac{1}{2\tau^2}\beta' K \beta\right).$$

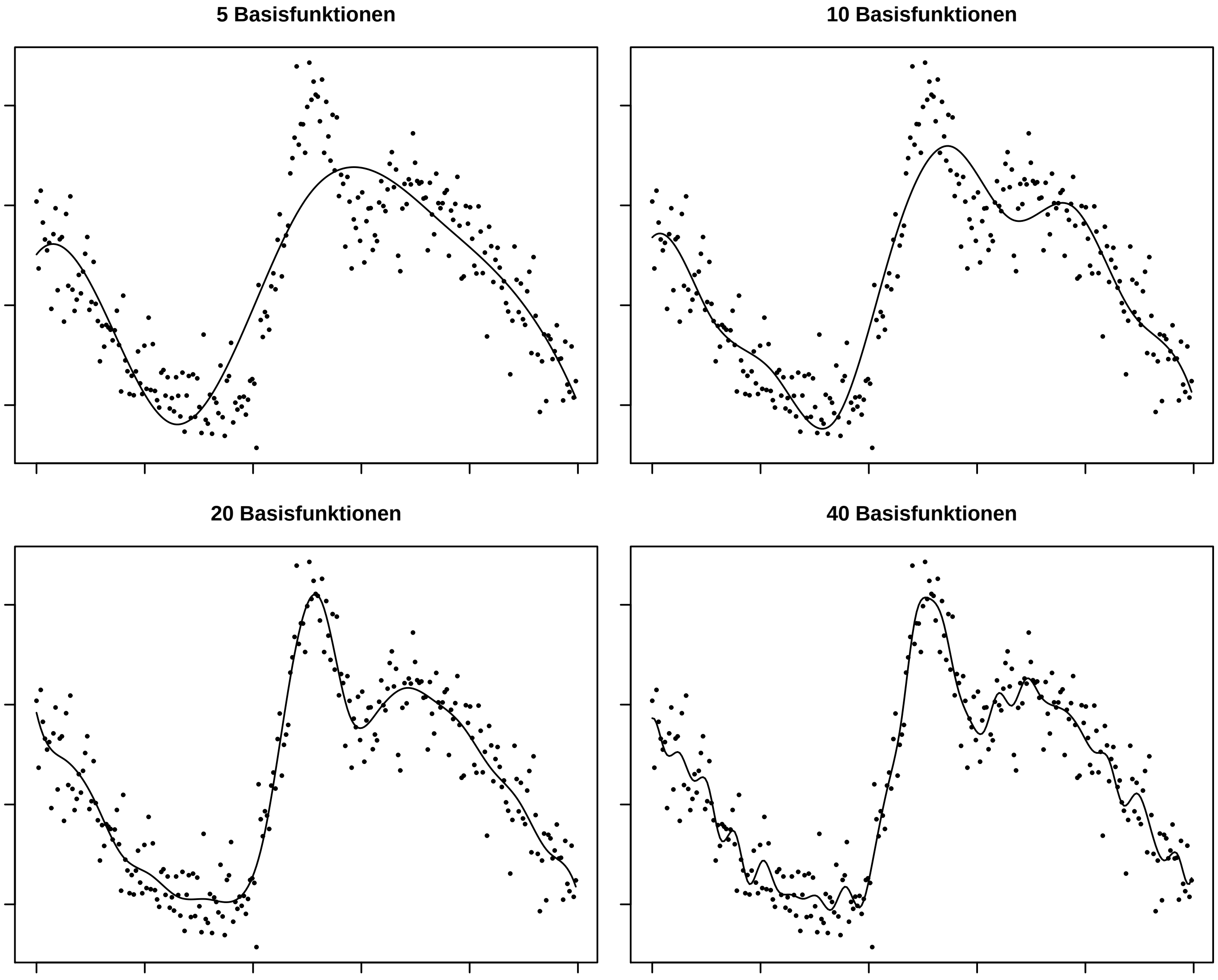
Penalisierte Splines

- Approximiere eine nichtparametrisch zu schätzende Funktion $f(x)$ durch eine Linearkombination von B-Spline Basisfunktionen $B_j(x)$

$$f(x) = \sum_j \beta_j B_j(x)$$



- B-Spline Schätzungen für variierende Anzahlen von Basisfunktionen:



- Unregularisierte Schätzungen hängen stark von der Anzahl der Basisfunktionen ab.

  $\Rightarrow$ Ergänze die Likelihood um einen Regularisierungs-Term der raue Funktionsschätzungen bestraft.

- Beliebter Ansatz: Bestrafung der quadrierten zweiten Ableitung

$$\text{pen}(f) = \lambda \int (f''(x))^2 dx.$$

- Einfache Approximation für B-Splines: Differenzen-Strafterme, z.B. für erste Differenzen

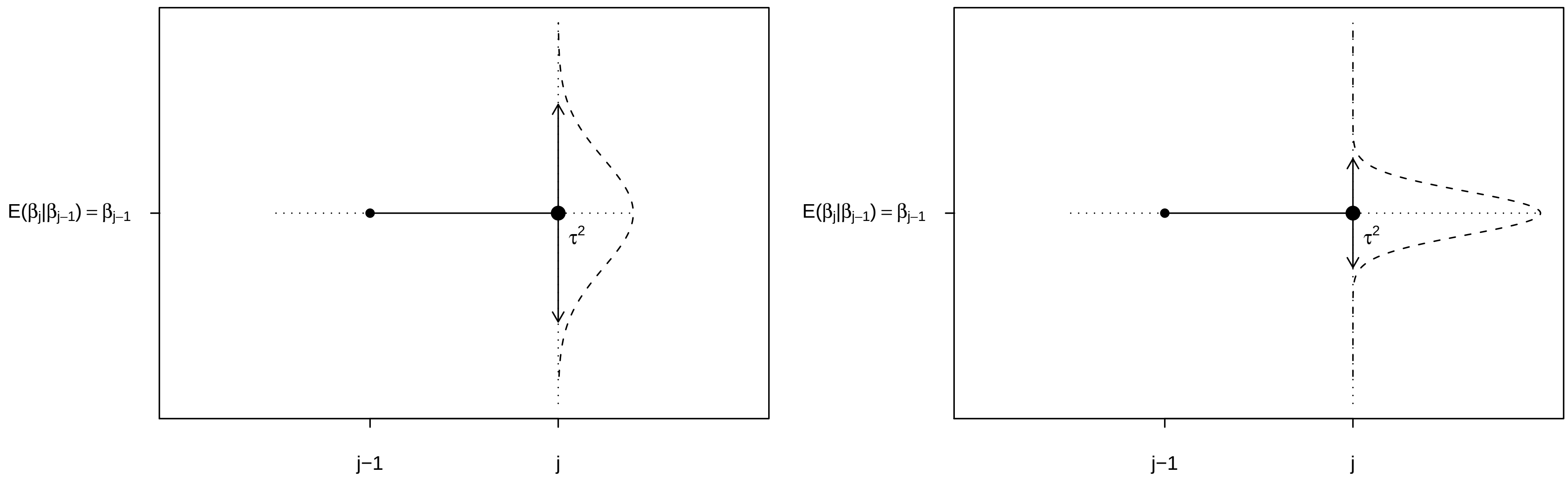
$$\text{pen}(\beta) = \lambda \sum_j (\beta_j - \beta_{j-1})^2 = \lambda \beta' K \beta$$

- Der Glättungsparameter λ bestimmt den Einfluss der Regularisierung auf die Schätzung

  $\Rightarrow$ Die Schätzung des Glättungsparameters ist die eigentliche Schwierigkeit.

- Random Walks liefern eine äquivalente bayesianische Formulierung der Differenzen-Regularisierung, z.B. durch Random Walks erster Ordnung

$$\beta_j = \beta_{j-1} + u_j, \qquad u_j \sim \mathsf{N}(0, \tau^2).$$



- Die gemeinsame Verteilung des Random Walks ergibt sich zu

$$p(\beta) \propto \exp\left(-\frac{1}{2\tau^2}\beta' K \beta\right).$$

- Man erhält die angekündigte quadratische Penalisierung bzw. multivariate Gauß-Priori.

- Aus der Bayesianischen Perspektive sind nichtparametrische Funktionen Realisationen stochastischer Prozesse

  $\Rightarrow$ Identifizierbarkeitsprobleme bei gleichzeitiger Schätzung von Trend und korrelierten Fehlern in longitudinalen Daten.

- Details in Fahrmeir & Kneib (2008): On the Identification of Trend and Correlation in Temporal and Spatial Regression.

Inferenz I: Gemischte Modelle

- Beobachtungsmodell:

$$\eta = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p.$$

- Multivariate Gauß-Priori für die Regressionskoeffizienten

$$p(\beta_j|\tau_j^2) \propto \exp\left(-\frac{1}{2\tau_j^2}\beta_j' K_j \beta_j\right).$$

- Entspricht einem Regressionsmodell mit zufälligen Effekten

$$\beta_j \sim N(0, \tau_j^2 K_j^-).$$

- Problem: Die Präzisionsmatrix K_j ist i.A. nicht invertierbar.

  $\Rightarrow$ Reparametrisierung um zu einem gemischten Modell mit eigentlichen zufälligen Effekten zu gelangen.

- Inferenz basierend auf gemischten Modellen:
  - Penalisierte Maximum Likelihood-Schätzung für die Regressionskoeffizienten.
  - Restricted Maximum Likelihood-Schätzung für die Varianzparameter.

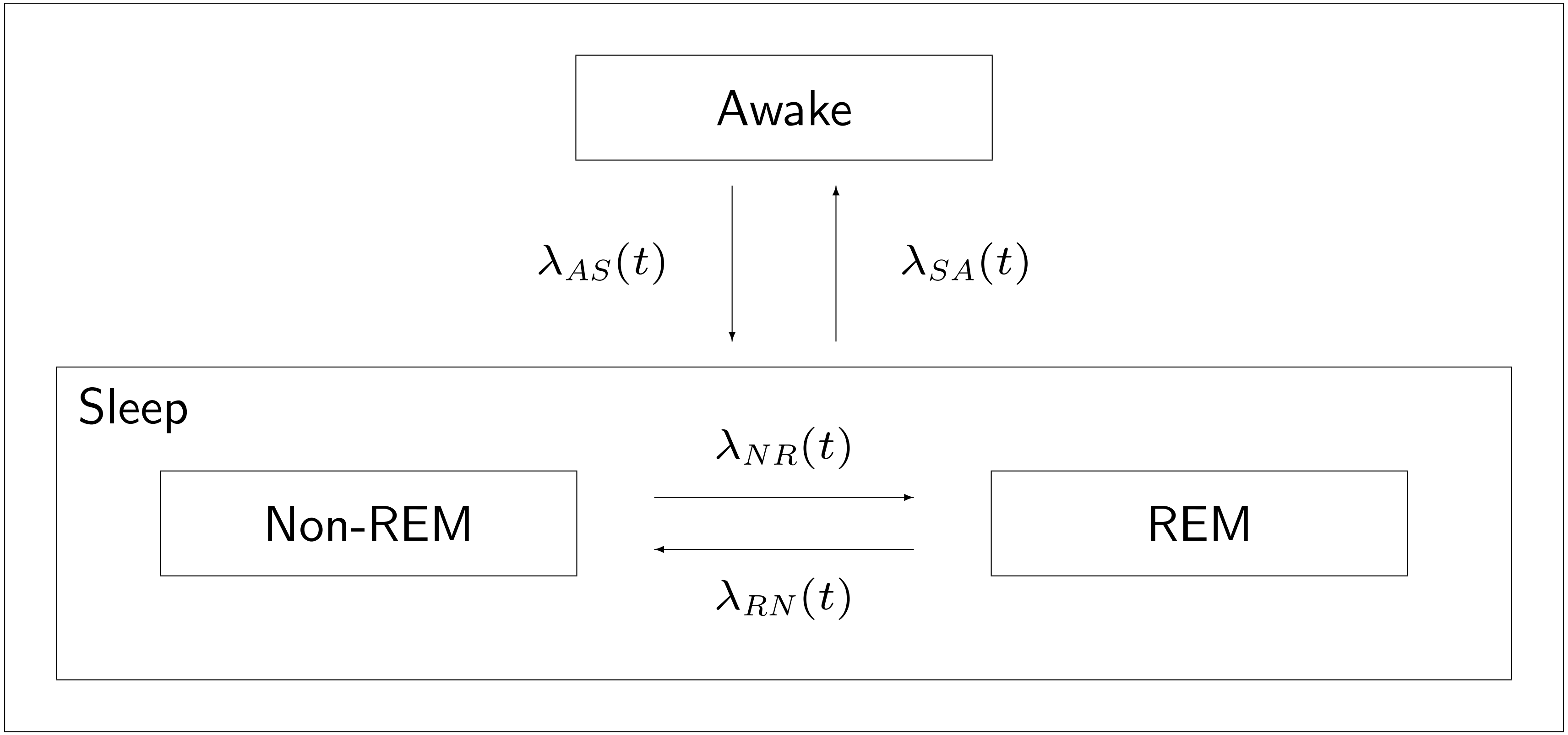
- Entspricht einem empirischen Bayes-Ansatz mit marginaler Likelihood-Schätzung für die Hyperparameter und Posteriori-Modus-Schätzung für die Regressionskoeffizienten.

- Das Prinzip ist einfach anwendbar für Modellterme mit Gauß-Prioris in denen die Präzisionsmatrix K_j nicht von weiteren Hyperparametern abhängt.

- Sehr flexibel in Hinblick auf die Verteilung der Zielvariable, da zur penalisierten ML-Schätzung nur leichte Modifikationen der üblichen ML-Schätzung nötig sind.

- Kneib & Hennerfeind (2008): Bayesian Semiparametric Multi-State Models.
  - Beschreibung von Übergangsintensitäten in Mehrstadien-Modellen in Analogie zu Hazardraten in Überlebenszeit-Modellen.
  - Simultane Schätzung der Baseline-Hazardraten (über penalisierte Splines) und der Kovariableneffekte (parametrisch oder zeitlich variierend).
  - Einbettung in Zählprozess-Theorie erlaubt die Likelihood-Konstruktion sowie die Modellüberprüfung über Martingal-Residuen.
  - Ebenfalls enthalten: Voller Bayes-Ansatz über Markov Chain Monte Carlo Simulationsverfahren.

- Beschreibung des menschlichen Schlafverlaufs:



- Kneib, Baumgartner & Steiner (2007): Semiparametric Multinomial Logit Models for Analysing Consumer Choice Behaviour.
  - Multinomiales Logit Modell zur Analyse des Markenwahl-Verhaltens von Konsumenten.
  - Nichtparametrische Modellierung der Effekte von Loyalität, Referenzpreis und Differenz zwischen Preis und Referenzpreis.
  - Modell-Überprüfung anhand von Scoring-Regeln.
- Kneib, Knauer & Küchenhoff (2007): A General Approach for Modelling Habitat Selection.
  - Multinomiales Logit Modell zur Analyse der Habitatwahl von Singvögeln und Braunbären.
  - Berücksichtigung der longitudinalen Struktur durch zufällige Effekte.
  - Offsets um Habitate unterschiedlicher Größe vergleichbar zu machen.

- Hofner, Kneib, Hartl & Küchenhoff (2008). Model Choice in Cox-Type Additive Hazard Regression Models with Time-Varying Effects.
  - Modellwahl in Hazard-Regressionsmodellen mit nichtparametrischen und zeitvariierenden Effekten.
  - Zweistufige Stepwise-Prozedur, die zwischen der Aufnahme von Kovariablen und verschiedenen Modellierungsformen unterscheidet.

Inferenz II: Markov Chain Monte Carlo Verfahren

- Beobachtungsmodell:

$$\eta = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p.$$

- Konditionale Gauß-Priori für die Regressionskoeffizienten:

$$p(\beta_j|\tau_j^2, \vartheta) \propto \exp\left(-\frac{1}{2\tau_j^2}\beta_j' K_j(\vartheta)\beta_j\right).$$

- Geeignete Hyperprioris für τ_j^2 und ϑ erlauben flexible Modellspezifikationen, die jedoch die Struktur des Updates der Regressionskoeffizienten unverändert lassen.

- Für normalverteilte Zielvariablen ergeben sich die vollständig bedingten Dichte in geschlossener Form und man erhält einen Gibbs-Sampler.

- Vollständig bedingte Verteilung für Regressionskoeffizienten β_j: $N(\mu_j, \Sigma_j)$ mit

$$\mu_j = \left(X_j'X_j + \frac{\sigma^2}{\tau_j^2}K_j\right)^{-1} X_j'(y - \eta + X_j\beta_j), \qquad \Sigma_j = \left(X_j'X_j + \frac{\sigma^2}{\tau_j^2}K_j\right)^{-1}$$

- Bei allgemeineren Zielvariablen Konstruktion von Vorschlagsdichten durch quadratische Approximation der vollständig bedingten Dichten (IWLS Proposal).

- Die modulare Struktur des MCMC-Algorithmus ist besonders geeignet bei hierarchisch formulierten Modellen.

- Im Folgenden am Beispiel der Regularisierung hochdimensionaler Kovariablenvektoren (wie im Mietspiegel-Beispiel).

- Ziel: Bevorzuge sparsame Modelle in denen eine große Zahl von Koeffizienten nahe oder gleich Null ist.

- Beispiel Ridge-Regression:

$$KQ_{\text{pen}}(\beta) = (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^{p} \beta_j^2 \to \min_{\beta}.$$

- Penalisierter KQ-Schätzer

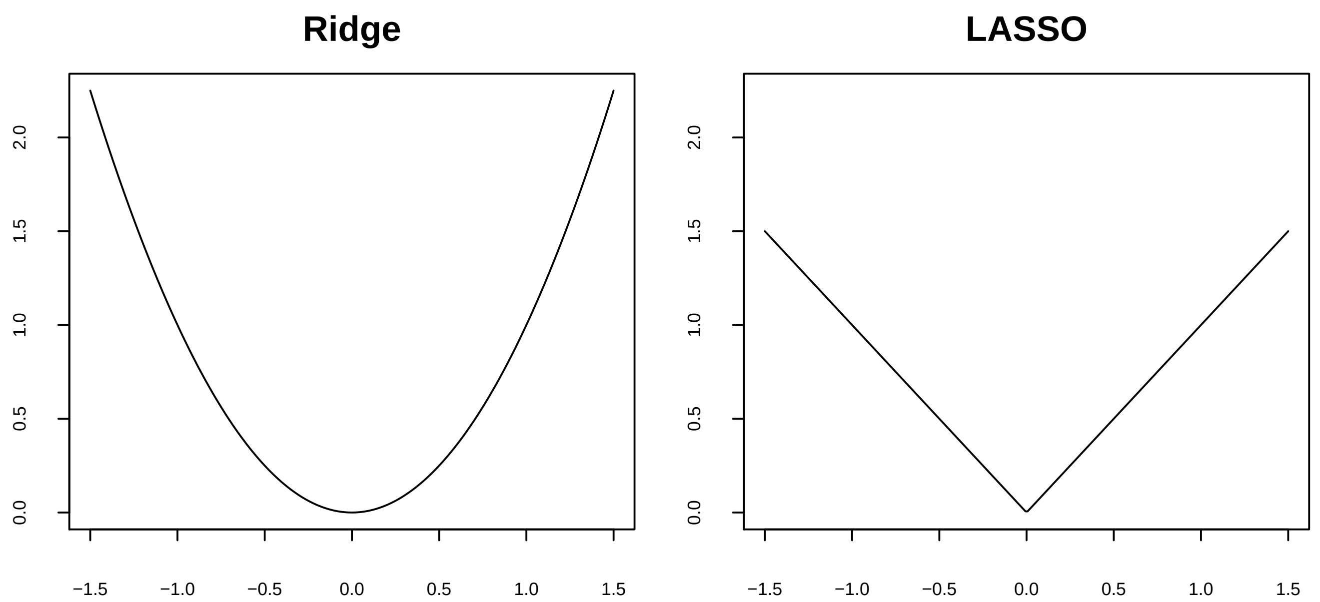
$$\hat{\beta} = (X'X + \lambda I)^{-1} X'y.$$

- Nachteil der Ridge-Regression: Die resultierenden Modelle sind nicht sparsam genug.

  $\Rightarrow$ Betrachte Strafterme mit Peak in der Null.

- Beispiel LASSO: Ersetze den quadratischen Strafterm durch den Absolutbetrag:

$$KQ_{\text{pen}}(\beta) = (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^{p} |\beta_j| \to \min_{\beta}.$$



- Äquivalenzbeziehungen zwischen Penalisierungen und Priori-Verteilungen:

Strafterm	**Priori-Dichte**	**Verteilung**
Ridge	$p(\beta_j) \propto \exp(-\lambda \beta_j^2)$	Gauß
LASSO	$p(\beta_j) \propto \exp(-\lambda \lvert\beta_j\rvert)$	Laplace
L_p	$p(\beta_j) \propto \exp(-\lambda \lvert\beta_j\rvert^p)$	Potenz-Exponential

- Die Laplace-Priori passt nicht zu unserer Standardannahme einer Gauß-Priori.

- Hierarchische Priori-Formulierung:

$$\lambda \longrightarrow \underset{\text{Lap}(\lambda)}{\beta} \qquad \text{vs.} \qquad \lambda \longrightarrow \underset{\text{Exp}(0.5\lambda^2)}{\tau^2} \longrightarrow \underset{\text{N}(0,\tau^2)}{\beta}$$

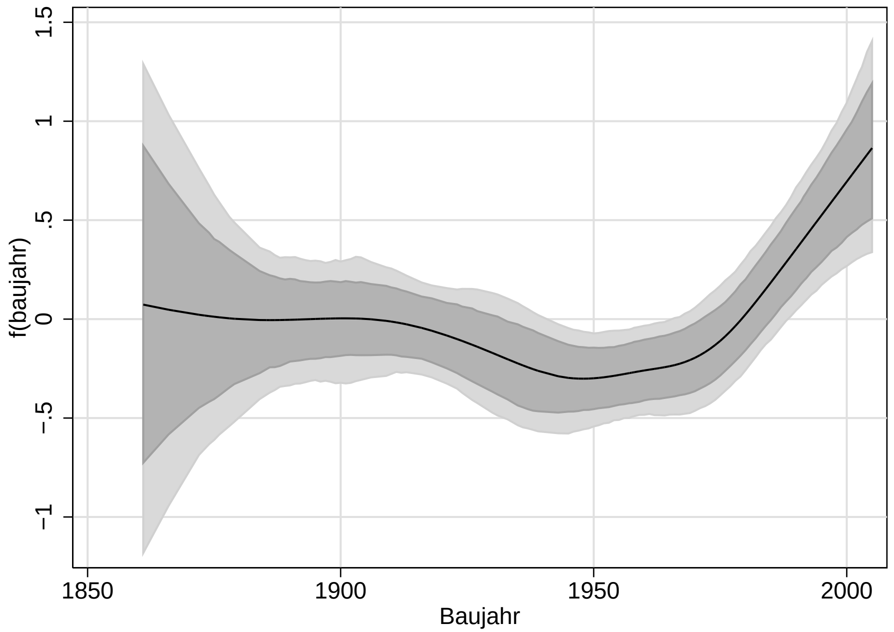
- Vorteil: Das Problem lässt sich auf die (einfacher zugängliche) Gauß-Formulierung zurückführen.

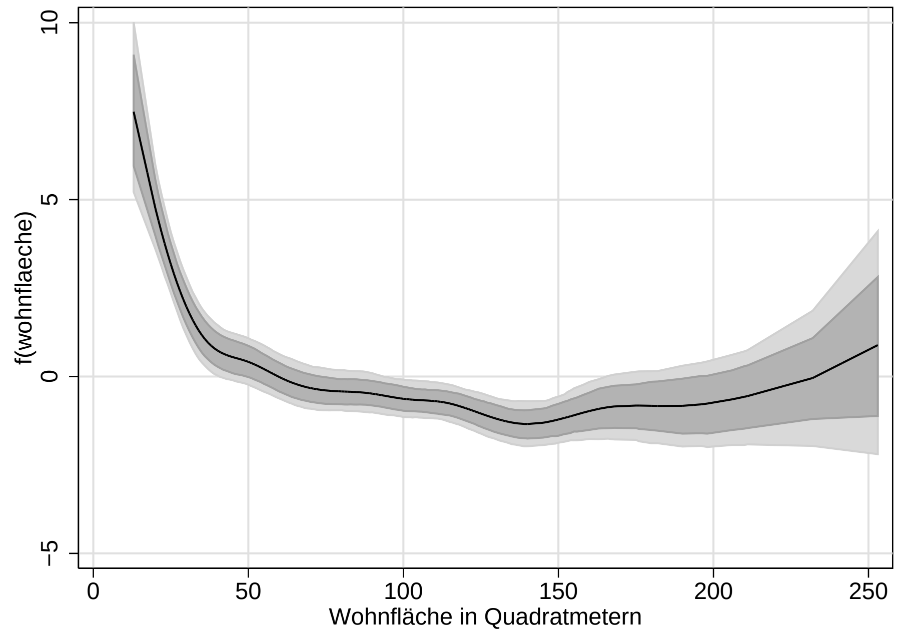
Münchner Mietspiegel: Ergebnisse

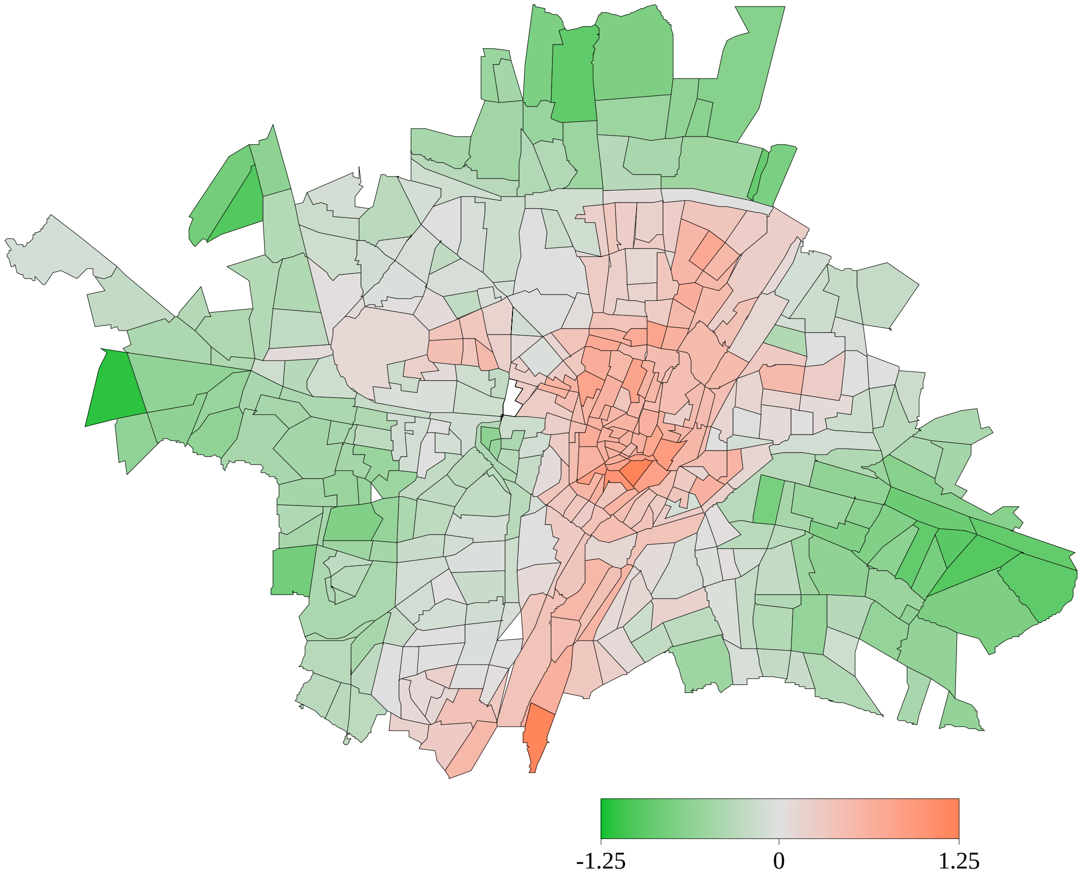
- Mietspiegel: Modellgleichung

$$nm = f_1(wohnflaeche) + f_2(baujahr) + f_3(bezirksviertel) + x'\beta + \varepsilon.$$

- Im Folgenden Ergebnisse bei LASSO-Regularisierung des Vektors β und Schätzung basierend auf MCMC.



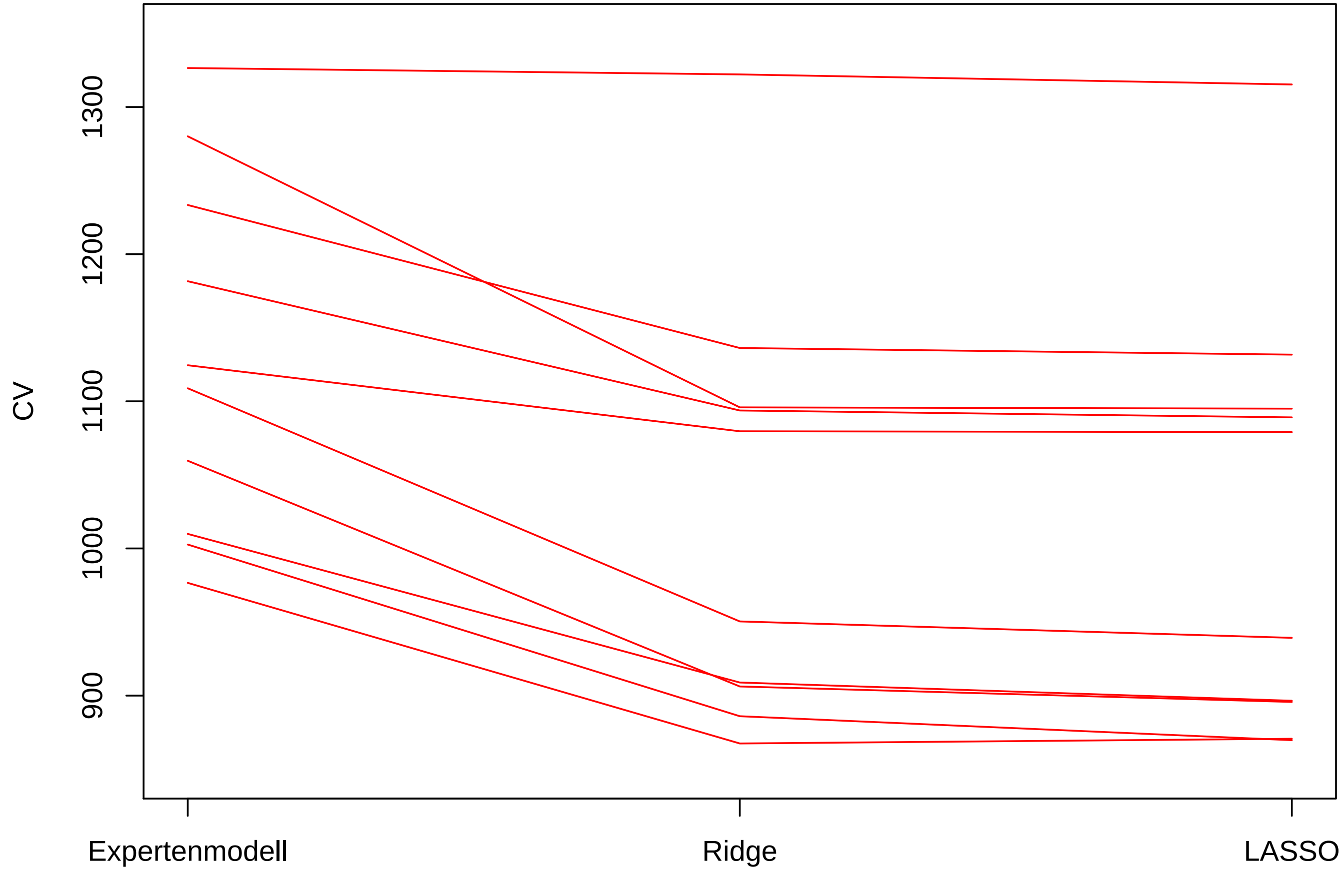




Räumlicher Effekt $f_3(s)$ in € pro Quadratmeter

- Interpretierbare Ergebnisse, aber was gewinnt man für die Prognose?

- Vergleich eines Expertenmodells (Subvektor + Transformation der Kovariablen), der Ridge-Regression und der LASSO-Regression über 10-fache Kreuzvalidierung.



$\Rightarrow$ Deutlich verbesserte Vorhersageeigenschaften durch Regularisierung!

Weitere Anwendungen

- Konrath, Fahrmeir & Kneib (2008). Bayesian Regularization in Structured Additive Regression Models for Survival Data.
  - Kombination von nichtparametrischer Regression und Regularisierung hochdimensionaler Kovariablenvektoren in Regressionsmodellen für Hazardraten.
  - Erweiterung der Regularisierung auf Modelle mit latenten Variablen, die zwischen einflussreichen und redundanten Kovariablen unterscheiden.
- Scheipl & Kneib (2008): Locally Adaptive Bayesian P-Splines with a Normal-Exponential-Gamma Prior.
  - Lokal adaptive Glättung mit bayesianischen P-Splines.
  - Ersetze den Gauß-Random Walk durch einen Random Walk mit Fehlern, die einer Verteilung mit Peak in Null und schweren Enden folgt.
  - Zusätzliche Adaptivität durch eine stückweise konstante Varianzfunktion.

- Kneib, Brezger & Crainiceanu (2008): Generalized Semiparametric Regression Models with Nonparametric Effects of Covariates Measured with Error.
  - Nichtparametrisch modellierte Effekte messfehlerbehafteter Kovariablen werden zu glatt geschätzt (Unterschätzung lokaler Extrema).
  - Korrektur durch Imputation der unbeobachteten wahren Kovariablenwerte im MCMC-Algorithmus.
  - Gegeben die imputierten Werte läuft die Schätzung dann wie zuvor, aber die Designmatrix des penalisierten Splines ändert sich in jeder Iteration

    $\Rightarrow$ Numerisch effiziente Umsetzung entscheidend.
  - Implementation funktioniert für beliebige Verteilungen der Zielvariablen und auch wenn mehrere Kovariablen fehlerbehaftet gemessen werden.

Inferenz III: Boosting

- Boosting ist ein flexibles funktionales Gradientenverfahren.

- Beschreibe Schätzprobleme durch eine Verlustfunktion $\rho(\eta, \text{Daten})$ (z.B., aber nicht notwendigerweise die negative Log-Likelihood).

- Iterative Schätzung geeigneter, einfacher Basis-Lernverfahren (Base-Learner) basierend auf dem negativen Gradienten der Verlustfunktion.

- Komponentenweises Boosting in Kombination mit frühzeitigem Abbrechen des Algorithmus erlaubt die Umsetzung von Variablenselektion und Modellwahl.

- Geeignete Base-Learner in strukturiert additiven Regressionsmodellen: Penalisierte Kleinste Quadrate Schätzer mit Prädiktionsmatrix

$$S_\lambda = X(X'X + \lambda K)^{-1} X'$$

(Kneib, Hothorn & Tutz (2009): Model Choice and Variable Selection in Geoadditive Regression Models).

- Entscheidend: Die Base-Learner müssen in ihrer Komplexität vergleichbar sein, um verzerrte Schätzungen zu vermeiden.

Zusammenfassung und Diskussion

- Einheitliche Formulierung einer flexiblen und reichhaltigen Klasse von Regressionsmodellen.

- Drei Inferenzkonzepte:
  - Gemischte Modelle (empirische Bayes-Schätzung).
  - Markov Chain Monte Carlo Simulationsverfahren (volle Bayes-Schätzung).
  - Boosting.

- Gemischte Modelle: Vorteile
  - Lässt sich sowohl frequentistisch als auch bayesianisch interpretieren.
  - Keine Abhängigkeit von Hyperprioris und keine Abhängigkeit von Sampling-Performance (im Gegensatz zu MCMC).
  - Auch nützlich in der theoretischen Untersuchung von strukturiert additiven Regressionsmodellen (Identifizierbarkeit, Normierbarkeit der Posteriori).

- Gemischte Modelle: Nachteile
  - Nicht modular.
  - Im Wesentlichen beschränkt auf Normalverteilungs-Prioris / quadratische Strafterme.
  - Unsicherheitsmaße und Tests basieren auf asymptotischen Argumenten.

- MCMC: Vorteile
  - Sehr flexibel durch modulare Behandlung der Modellschätzung.
  - Nutzt die hierarchische Modellformulierung aus.
  - Linear wachsende Komplexität.
  - Exakte Unsicherheitsmaße auch für komplexe Funktionen der Parameter.
- MCMC: Nachteile
  - Setzt die volle Spezifikation einer Likelihood voraus.
  - Sensitivität in Bezug auf Hyperprioris und Hyperparameter.
  - Benötigt (in manchen Fällen) das Tunen von Vorschlagsdichten.
  - Konvergenz und Mixing müssen untersucht werden.
  - Bei Verwendung (teilweise) uneigentlicher Prioris muss die Normierbarkeit der Posteriori gesichert sein (Fahrmeir & Kneib (2009): Propriety of Posteriors in Structured Additive Regression Models: Theory and Empirical Evidence).

- Boosting: Vorteile
  - Eingebaute Variablenselektion und Modellwahl.
  - Flexible Formulierung des Schätzproblems über eine Verlustfunktion (robuste Schätzung, Quasi-Likelihood).
  - Base-learner ergeben eine ähnliche Modularität wie bei MCMC.

- Boosting: Nachteile
  - Nur Punktschätzer, Unsicherheitsmaße und Tests schwierig.

Bestandteile der Habilitationsschrift

- Kneib, T., Baumgartner, B. & Steiner, W. J. (2007). Semiparametric Multinomial Logit Models for Analysing Consumer Choice Behaviour. AStA Advances in Statistical Analysis, 91, 225–244.
- Kneib, T., Knauer, F. & Küchenhoff, H. (2007). A General Approach for Modelling Habitat Selection. In Revision bei Environmental and Ecological Statistics.
- Fahrmeir, L. & Kneib, T. (2009). Propriety of Posteriors in Structured Additive Regression Models: Theory and Empirical Evidence. Journal of Statistical Planning and Inference, 139, 843–859.
- Fahrmeir, L. & Kneib, T. (2008). On the Identification of Trend and Correlation in Temporal and Spatial Regression. In: Shalab & Heumann, C. (Hrsg.): Recent advances in linear models and related areas, Springer.
- Kneib, T., Hothorn, T. & Tutz, G. (2009). Model Choice and Variable Selection in Geoadditive Regression Models. Erscheint in Biometrics.
- Scheipl, F. & Kneib, T. (2008). Locally Adaptive Bayesian P-Splines with a Normal-Exponential-Gamma Prior. In Revision bei Computational Statistics & Data Analysis.
- Hofner, B., Kneib, T., Hartl, W. & Küchenhoff, H. (2008). Model Choice in Cox-Type Additive Hazard Regression Models with Time-Varying Effects. Technical Report.
- Konrath, S., Fahrmeir, L. & Kneib, T. (2008). Bayesian Regularization in Structured Additive Regression Models for Survival Data. Technical Report.
- Kneib, T., Brezger, A. & Crainiceanu, C. M. (2008). Generalized Semiparametric Regression Models with Nonparametric Effects of Covariates Measured with Error. Technical Report.
- Kneib, T. & Hennerfeind, A. (2008). Bayesian Semiparametric Multi-State Models. Statistical Modelling, 8, 169–198.

Danke

Nicole Augustin – Thomas Augustin – Bernhard Baumgartner – Christiane Belitz – Axel Benner – Jan Beyersmann – Susanne Breitner – Andreas Brezger – Carmen Cadarso-Suárez – Ciprian M. Crainiceanu – Christiane Dargatz – Ludwig Fahrmeir – Nora Fenske – Sonja Greven – Wolfgang Hartl – Susanne Heim – Leonhard Held – Andrea Hennerfeind – Benjamin Hofner – Michael Höhle – Hajo Holzmann – Ludwig Hothorn – Torsten Hothorn – Stefanie Kalus – Göran Kauermann – Stephan Klasen – Carmen Klement – Felix Knauer – Barbara Kneib – Gottfried Kneib – Susanne Konrath – Tatyana Krivobokova – Daniela Kropf – Helmut Küchenhoff – Stefan Lang – Fritz Leisch – Ulrich Mansmann – Brian D. Marx – Jörg Müller – Axel Munk – Thomas Petzoldt – Jan Priebe – Håvard Rue – David Ruppert – Fabian Scheipl – Matthias Schmid – Martin Schumacher – Wilfried Seidl – Stefan Sperlich – Ulrich Stadtmüller – Winfried J. Steiner – Alexander Strasak – Carolin Strobl – Viola Svejdar – Gerhard Tutz – Alexander Yassouridis – Achim Zeileis – Walter Zucchini