Einführung in INLA

Integrated Nested Laplace Approximation

Martje Rave

Motivation: Warum INLA?

- In der Bayes-Statistik benötigen wir oft Posterior-Verteilungen.
- Für viele Modelle (z. B. GLMMs, räumlich-zeitliche Modelle) ist die exakte Berechnung schwer.
- Klassische Lösung: MCMC — aber kann teuer, langsam, schwer zu diagnostizieren sein.
- INLA bietet eine schnelle und präzise Approximation für bestimmte Modelle: latente Gausssche Modelle.

Grundidee von INLA

Ziel

Gegeben Beobachtungen y, möchten wir Posterior-Marginalverteilungen

$$p(\theta_j|y) \quad \text{und} \quad p(\tau_k|y)$$

für Modellparameter θ und Hyperparameter τ berechnen.

- INLA verwendet die Laplace-Approximation zur Berechnung dieser Posteriori.
- Dabei wird die bedingte Dichte $p(\theta|\tau, y)$ durch eine Normalverteilung approximiert.

Hierarchisches Modell in INLA

Ein typisches Modell hat die Struktur:

$$
\begin{aligned}
y &\sim F(\theta), \\
\theta_j &\sim \mathcal{N}(\mu_j, \tau_j^{-1}), \\
\tau_j &\sim \text{Prior}(\cdot)
\end{aligned}
$$

- y: Beobachtungen
- θ: latente (Gaußsche) Felder
- τ: Hyperparameter (z.B. Varianzen, Glättungsparameter)

Laplace-Approximation – Grundidee

Gesucht:

$$\int_{-\infty}^{\infty} \exp(-nh(x))dx$$

- Idee: Entwickle $h(x)$ um das Minimum $\tilde{x}$:

$$h(x) \approx h(\tilde{x}) + \frac{1}{2}h''(\tilde{x})(x-\tilde{x})^2$$

- Dann:

$$\int \exp(-nh(x))dx \approx \exp(-nh(\tilde{x}))\sqrt{\frac{2\pi}{nh''(\tilde{x})}}$$

Laplace-Approximation – Grundidee

Gesucht:

$$\int_{-\infty}^{\infty} \exp(-nh(x))dx$$

- Idee: Entwickle $h(x)$ um das Minimum $\tilde{x}$:

$$h(x) \approx h(\tilde{x}) + \frac{1}{2}h''(\tilde{x})(x-\tilde{x})^2$$

- Dann:

$$\int \exp(-nh(x))dx \approx \exp(-nh(\tilde{x}))\sqrt{\frac{2\pi}{nh''(\tilde{x})}}$$

Anwendung: Berechnung von Posterior-Dichten durch Approximation von Integralen.

Marginale Posteriori in INLA

1. **Schritt:** Approximiere $p(\theta|\tau, y)$ durch Laplace.
2. **Schritt:** Berechne die marginale Posteriori von τ über:

$$p(\tau|y) \approx \frac{p(\theta^*(\tau), \tau, y)}{p(\theta^*|\tau, y)}$$

3. **Schritt:** Verwende numerische Integration zur Approximation von:

$$p(\theta_j|y) = \int p(\theta_j|\tau, y)p(\tau|y)d\tau$$

Vorteil: Alle Schritte sind deterministisch und relativ schnell!

Beispiel: Laplace-Approximation im Poisson-Gamma-Modell

- Simulierte Daten aus Poisson-Modell mit Jeffreys-Prior.
- Vergleich zwischen exaktem Posterior-Erwartungswert und Laplace-Approximation:

n	Erwartungswert	Approximation	Fehler
1	27.50	25.70	1.80
10	25.45	25.25	0.20
100	26.18	26.15	0.02

Fazit: Für große n ist die Approximation sehr genau.

Vorteile und Grenzen von INLA

Vorteile:

- Sehr schnell für große Datensätze
- Hohe Genauigkeit im Vergleich zu MCMC
- Gut geeignet für latente Gaußsche Modelle

Vorteile und Grenzen von INLA

Vorteile:

- Sehr schnell für große Datensätze
- Hohe Genauigkeit im Vergleich zu MCMC
- Gut geeignet für latente Gaußsche Modelle

Grenzen:

- Eingeschränkt auf bestimmte Modellklassen
- Laplace-Approximation kann in sehr komplexen Modellen ungenau sein

Zusammenfassung

- INLA ist eine effiziente Methode zur Approximation bayesscher Posteriori.
- Verwendet deterministische Laplace-Näherungen statt stochastischer MCMC-Sampling.
- Besonders geeignet für GLMMs, Zeitreihen- und räumliche Modelle.
- In R verfügbar über das Paket `INLA`.

Nächste Sitzung: Anwendung von INLA auf reale Daten in R.