Bayesian Learning - Lecture 9

Mattias Villani

Division of Statistics
Department of Computer and Information Science
Linköping University

Lecture overview

- Metropolis-Hastings within Gibbs
- Heteroscedastic regression
- Regularized regression with Gibbs
- RStan demo

Metropolis-Hastings within Gibbs

- **Gibbs sampling** from $p(\theta_1, \theta_2, \theta_3|\mathbf{y})$
  - Sample $p(\theta_1|\theta_2, \theta_3, \mathbf{y})$
  - Sample $p(\theta_2|\theta_1, \theta_3, \mathbf{y})$
  - Sample $p(\theta_3|\theta_1, \theta_2, \mathbf{y})$
- When a **full conditional is not easily sampled** we can simulate from it using MH.
- Example: at ith iteration, propose θ_2 from $q(\theta_2|\theta_1, \theta_3, \theta_2^{(i-1)}, \mathbf{y})$. Accept/reject.
- Gibbs sampling is a special case of MH when $q(\theta_2|\theta_1, \theta_3, \theta_2^{(i-1)}, \mathbf{y}) = p(\theta_2|\theta_1, \theta_3, \mathbf{y})$, which gives $\alpha = 1$. Always accept.

Heteroscedastic regression

- **Heteroscedastic regression**:

$$y_i = x_i'\beta + \varepsilon_i$$

where the errors are heteroscedastic

$$\varepsilon_i \sim N\left[0, \sigma^2 \exp\left(x_i'\gamma\right)\right].$$

- Priors: Multivariate normal for β and γ. Inv-χ^2 for σ^2.
- **Gibbs sampling**:
  - $\beta, \sigma^2 | \gamma, \mathbf{y}, X$
  - $\gamma | \beta, \sigma^2, \mathbf{y}, \mathbf{X}$

Heteroscedastic regression

- Draws from $\beta, \sigma^2 | \gamma, \mathbf{y}, X$ can be easily obtained since we can rewrite the model as

$$\tilde{y}_i = \tilde{x}_i' \beta + \tilde{\varepsilon}_i$$

  where

  - $\tilde{y}_i = \exp\left(-x_i'\gamma/2\right) y_i$
  - $\tilde{x}_i' = \exp\left(-x_i'\gamma/2\right) x_i'$
  - $\tilde{\varepsilon}_i = \exp\left(-x_i'\gamma/2\right) \varepsilon_i$.
  - Note that $Var(\tilde{\varepsilon}_i) = \sigma^2$, so **homoscedastic**.

- $p(\gamma | \beta, \sigma^2, \mathbf{y}, \mathbf{X})$ is non-standard, but we can use **MH to sample** from proposal $N(\hat{\delta}, \Sigma)$

- $\hat{\delta}$ and Σ can be obtained from approximate maximum likelihood estimators. (but there are better proposals).

Regularized regression with Gibbs

- Recap: The joint posterior of β, σ^2 and λ is

$$
\begin{aligned}
\beta|\sigma^2,\lambda,\mathbf{y},\mathbf{X} &\sim N\left(\mu_n,\Omega_n^{-1}\right) \\
\sigma^2|\lambda,\mathbf{y},\mathbf{X} &\sim Inv-\chi^2\left(\nu_n,\sigma_n^2\right) \\
p(\lambda|\mathbf{y},\mathbf{X}) &\propto \sqrt{\frac{|\Omega_0|}{|X'X+\Omega_0|}}\left(\frac{\nu_n\sigma_n^2}{2}\right)^{-\nu_n/2}\cdot p(\lambda)
\end{aligned}
$$

  where $p(\lambda)$ is the $Inv-\chi^2$ prior for λ.
- This is the **conditional-marginal decomposition**

$$p(\beta,\sigma^2,\lambda|\mathbf{y},\mathbf{X}) = p(\beta|\sigma^2,\lambda,\mathbf{y},\mathbf{X})p(\sigma^2|\lambda,\mathbf{y},\mathbf{X})p(\lambda|\mathbf{y},\mathbf{X})$$

- **Gibbs sampling** can instead be used:
  - Sample $\beta|\sigma^2,\lambda,\mathbf{y},\mathbf{X}$ from Normal
  - Sample $\sigma^2|\beta,\lambda,\mathbf{y},\mathbf{X}$ from Inv-χ^2
  - Sample $\lambda|\beta,\sigma^2,\mathbf{y},\mathbf{X}$ from Inv-χ^2
- Note that λ is now **easy** to simulate **once we condition** on β and σ^2.