Inference in Ising Models

Bhaswar B. Bhattacharya

(Joint work with Sumit Mukherjee)

Stanford University

11th December, 2015
The Ising Model

- The data is a vector of dependent ±1 random variables

\[ \sigma = (\sigma_1, \sigma_2, \cdots, \sigma_N). \]

- The dependence between the coordinates of \( \sigma \) are modeled by a one-parameter exponential family on \( S_N := \{-1, 1\}^N \)

\[ \mathbb{P}_\beta(\sigma = \tau) = 2^{-N} \exp \left\{ \frac{1}{2} \beta H_N(\tau) - F_N(\beta) \right\}. \]

- Here
  - \( \beta > 0 \) is the *natural parameter (inverse temperature)*,
  - the sufficient statistic is a quadratic form:

\[ H_N(\tau) = \tau' J_N \tau = \sum_{1 \leq i, j \leq N} J_N(i, j) \tau_i \tau_j \]

for a symmetric matrix \( J_N \) with zeros on the diagonals.

- \( F_N(\beta) \) is the *log-normalizing constant* which is determined by the condition \( \sum_{\tau \in S_N} \mathbb{P}\{\sigma = \tau\} = 1. \)
Goal: We want to estimate the parameter $\beta$ given one realization from the model.
Goal: We want to estimate the parameter $\beta$ given one realization from the model.

Maximum Likelihood Estimation: Very difficult due to appearance of an intractable normalizing constant $F_N(\beta)$ in the likelihood.
The Problem

- **Goal**: We want to estimate the parameter $\beta$ given one realization from the model.

- **Maximum Likelihood Estimation**: Very difficult due to appearance of an intractable normalizing constant $F_N(\beta)$ in the likelihood.

- **Maximum Pseudo-Likelihood Estimation (MPLE)**: The pseudo-likelihood is obtained by multiplying all of the conditional likelihoods (Besag 1974).

  - The distribution of $\sigma_i$ given $\{\sigma_j, j \neq i\}$ takes the values $+1$ and $-1$ with probabilities $\frac{e^{\beta m_i}}{e^{\beta m_i} + e^{-\beta m_i}}$ and $\frac{e^{-\beta m_i}}{e^{\beta m_i} + e^{-\beta m_i}}$ respectively, where $m_i = \sum_{j=1}^{n} J_N(i, j) \sigma_j$.

  - The value of $\beta$ which maximizes the pseudo-likelihood is the pseudo-likelihood estimator $\hat{\beta}_N$. 
Consistency of the MPLE

Theorem (Chatterjee (2007))

If \( \sup_N \|J_N\| < \infty \) and

\[
\lim_{N \to \infty} \inf \frac{1}{N} F_N(\beta_0) > 0,
\]

then the MPLE \( \hat{\beta}_N \) is \( \sqrt{N} \)-consistent at \( \beta = \beta_0 \).
Theorem (Chatterjee (2007))

If \( \sup_N \|J_N\| < \infty \) and

\[
\lim \inf_{N \to \infty} \frac{1}{N} F_N(\beta_0) > 0,
\]

then the MPLE \( \hat{\beta}_N \) is \( \sqrt{N} \)-consistent at \( \beta = \beta_0 \).
Consistency of the MPLE: Our Results

Theorem (B.-Mukherjee (2015))

Let \( \sup_{N \geq 1} ||J_N|| < \infty \), and \( \beta_0 > 0 \) be fixed. Suppose \( \{a_N\}_{N \geq 1} \) is a sequence of positive reals diverging to \( \infty \) such that

\[
0 < \lim_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{a_N} F_N(\beta_0 - \delta) \leq \lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{a_N} F_N(\beta_0 + \delta) < \infty.
\]

Then (under technical conditions) the MPLE \( \hat{\beta}_N \) is \( \sqrt{a_N} \)-consistent for \( \beta = \beta_0 \).
Consistency of the MPLE: Our Results

Theorem (B.-Mukherjee (2015))

Let \( \sup_{N \geq 1} ||J_N|| < \infty \), and \( \beta_0 > 0 \) be fixed. Suppose \( \{a_N\}_{N \geq 1} \) is a sequence of positive reals diverging to \( \infty \) such that

\[
0 < \lim_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{a_N} F_N(\beta_0 - \delta) \leq \lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{a_N} F_N(\beta_0 + \delta) < \infty.
\]

Then (under technical conditions) the MPLE \( \hat{\beta}_N \) is \( \sqrt{a_N} \)-consistent for \( \beta = \beta_0 \).

- (Erdős-Renyi Graphs) Suppose \( G_N \sim \mathcal{G}(N, p_N) \) with \( p_N \gg \frac{\log N}{N} \).
  Then \( \hat{\beta}_N \) is \( \sqrt{\frac{1}{p_N}} \) consistent if \( \beta < 1 \), and \( \sqrt{N} \) consistent if \( \beta > 1 \).
Theorem (B.-Mukherjee (2015))

Let \( \sup_{N \geq 1} \| J_N \| < \infty \), and \( \beta_0 > 0 \) be fixed. Suppose \( \{a_N\}_{N \geq 1} \) is a sequence of positive reals diverging to \( \infty \) such that

\[
0 < \lim_{\delta \to 0} \lim_{N \to \infty} \frac{1}{a_N} F_N (\beta_0 - \delta) \leq \lim_{\delta \to 0} \lim_{N \to \infty} \frac{1}{a_N} F_N (\beta_0 + \delta) < \infty.
\]

Then (under technical conditions) the MPLE \( \hat{\beta}_N \) is \( \sqrt{a_N} \)-consistent for \( \beta = \beta_0 \).

- (Erdős-Renyi Graphs) Suppose \( G_N \sim \mathcal{G}(N, p_N) \) with \( p_N \gg \frac{\log N}{N} \).
  Then \( \hat{\beta}_N \) is \( \sqrt{\frac{1}{p_N}} \) consistent if \( \beta < 1 \), and \( \sqrt{N} \) consistent if \( \beta > 1 \).

- (Regular Graphs) Suppose \( G_N \) is a sequence of \( d_N \)-regular graphs. If \( d_N \to \infty \), then \( \hat{\beta}_N \) is \( \sqrt{\frac{N}{d_N}} \) consistent if \( \beta < 1 \), and \( \sqrt{N} \) consistent if \( \beta > 1 \).
Figure: The MPLE and the 1-standard deviation error bar in an Ising model on \( G_N \sim \mathcal{G}(N, p(N)) \) with \( N = 2000 \) and \( p(N) = N^{-\frac{1}{3}} \), averaged over 100 repetitions for a sequence of values of \( \beta \in [0, 2] \).
When Consistent Estimation is Impossible?

Theorem (B.-Mukherjee (2015))

Suppose the log-normalizing constant $F_N(\beta_0) = O(1)$. Then there does not exist any consistent sequence of estimators in the interval $[0, \beta_0]$. 

$\text{Dense Graphs}$ Suppose $G_N$ is a sequence of dense graphs converging a graphon $W$ with maximum eigenvalue of $\lambda_1(W)$ (Borgs et al. (2008)). Then $\hat{\beta}_N$ is inconsistent for $\beta < \frac{1}{\lambda_1(W)}$, and $\sqrt{N}$ consistent for $\beta > 1$. Moreover, there exists no sequence of consistent estimators for $\beta < \frac{1}{\lambda_1(W)}$, where $\lambda_1(W)$ denotes the largest eigenvalue of $W$. 

Bhaswar B. Bhattacharya

Inference in Ising Models 6 / 9
When Consistent Estimation is Impossible?

**Theorem (B.-Mukherjee (2015))**

Suppose the log-normalizing constant $F_N(\beta_0) = O(1)$. Then there does not exist any consistent sequence of estimators in the interval $[0, \beta_0]$.

- **(Dense Graphs)** Suppose $G_N$ is a sequence of dense graphs converging a graphon $W$ with maximum eigenvalue of $\lambda_1(W)$ (Borgs et al. (2008)).
  - Then $\hat{\beta}_N$ is **inconsistent** for $\beta < \frac{1}{\lambda_1(W)}$, and $\sqrt{N}$ consistent for $\beta > 1$.
  - Moreover, there exists **no sequence of consistent estimators** for $\beta < \frac{1}{\lambda_1(W)}$, where $\lambda_1(W)$ denotes the largest eigenvalue of $W$. 
Example: Erdős-Rényi Graphs

Figure: The power of the MP-test for the Ising model on an Erdős-Rényi random graph $G(N, p)$ as a function of $p$ and $\beta$, with $N = 500$. In this case, the limiting graphon is $W \equiv p$, and $\lambda_1(W) = p$. Note the phase transition curve $\beta(p) = \frac{1}{p}$ above which the MP-test has power 1.
US Presidential Elections Data

- The dataset consists of political colors of the 48 states in the continental US (excluding Alaska, Hawaii, and Washington D.C.) in the last 26 presidential elections held during 1912-2012.
- Each state is assigned 1 for Democratic (colored blue) and -1 for Republican (colored red).
- The vertices of the *neighborhood graph* are the states and there is an edge between two states if they share a border.

1924 Elections: $\hat{\beta} = 2.84$  
1992 Elections: $\hat{\beta} = 0.37$  
2012 Elections: $\hat{\beta} = 0.96$
Thank You