Estimation of a Ramsay-Curve IRT Model using the Metropolis-Hastings Robbins-Monro Algorithm

Scott Monroe & Li Cai

IMPS 2012, Lincoln, Nebraska

UCLA
Outline

1. Introduction and Motivation
2. Review of Ramsay Curves
3. Review of the MH-RM Algorithm
4. Simulation Study
5. Empirical Application
6. Conclusion
1 Introduction and Motivation
2 Review of Ramsay Curves
3 Review of the MH-RM Algorithm
4 Simulation Study
5 Empirical Application
6 Conclusion
The need to estimate the latent trait distribution has received increasing attention in recent years.


However, the EM algorithm does not yield the observed information matrix upon convergence. Hence, no standard errors are available.
The need to estimate the latent trait distribution has received increasing attention in recent years.


However, the EM algorithm does not yield the observed information matrix upon convergence. Hence, no standard errors are available.

Outline

1. Introduction and Motivation
2. Review of Ramsay Curves
3. Review of the MH-RM Algorithm
4. Simulation Study
5. Empirical Application
6. Conclusion
Review of Ramsay Curves

For full details, see Woods and Thissen (2006). Given “observed” latent trait scores, RCs can estimate the distribution’s shape.

Essentially, the method uses B-spline regression, with a small number of parameters.
Outline

1. Introduction and Motivation
2. Review of Ramsay Curves
3. Review of the MH-RM Algorithm
4. Simulation Study
5. Empirical Application
6. Conclusion
Metropolis-Hastings Robbins-Monro (Cai, 2010) is motivated by:

Fisher (1925)

The gradient of the observed log likelihood

\[ = \]

the expectation of the gradient of the complete log likelihood.
A Review of MH-RM

Metropolis-Hastings Robbins-Monro (Cai, 2010) is motivated by:

Fisher (1925)

The gradient of the observed log likelihood

\[
\text{the expectation of the gradient of the complete log likelihood.}
\]

Iterative approach:

- Use MH sampler to find an MC approximation of the expectation.
  - Use this approximation to update the parameter estimates.
Metropolis-Hastings Robbins-Monro (Cai, 2010) is motivated by:

Fisher (1925)

The gradient of the observed log likelihood

\[ \text{the expectation of the gradient of the complete log likelihood.} \]

Iterative approach:

- Use MH sampler to find an MC approximation of the expectation.
  - Use this approximation to update the parameter estimates.
- Since the approximations are noisy, use the RM method to filter.
  - The filtering operates through decreasing gain constants.
Stage 1 (iterations 1-800): move to “neighborhood” of MLE.
MH-RM Estimation in a Picture

- **Stage 1** (iterations 1-800): move to “neighborhood” of MLE.
- **Stage 2** (iterations 801-1000): find starting values for Stage 3 by averaging estimates across Stage 2 iterations.

![Graph showing parameter estimates over iterations](image-url)
MH-RM Estimation in a Picture

- **Stage 1** (iterations 1-800): move to “neighborhood” of MLE.
- **Stage 2** (iterations 801-1000): find starting values for Stage 3 by averaging estimates across Stage 2 iterations.
- **Stage 3** (iterations 1001-convergence): obtain MLE by using decreasing gain constants to average out the noise.

![Graph showing parameter estimates over iterations](image-url)
The observed information can be expressed as a function of the complete log likelihood, involving its gradient and hessian.

With MH-RM, the pieces used in the approximation can be calculated by:

- recursive approximation, as in Cai (2010); or
- Monte Carlo approximation, as in Diebolt and Ip (1996).
1. Introduction and Motivation
2. Review of Ramsay Curves
3. Review of the MH-RM Algorithm
4. Simulation Study
5. Empirical Application
6. Conclusion
## Simulation Study

### Simulation Study Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>1000</td>
</tr>
<tr>
<td>Number of Items</td>
<td>25</td>
</tr>
<tr>
<td>Slope Dist.</td>
<td>$\mathcal{N}(1.7, 0.3)$</td>
</tr>
<tr>
<td>Location Dist.</td>
<td>$\mathcal{N}(0, 1)$, truncated at $\pm 2$</td>
</tr>
<tr>
<td>True $g(\theta</td>
<td>\eta)$</td>
</tr>
<tr>
<td>Replications</td>
<td>100</td>
</tr>
<tr>
<td>Estimation Methods</td>
<td>MH-RM and BA-EM</td>
</tr>
</tbody>
</table>
The Skewed and Bimodal densities were found as mixtures of normals.

Latent trait scores were drawn by rejection sampling.

Item responses were simulated in accordance with the 2PL model.
Both MH-RM and EM perform maximum marginal likelihood estimation, so we shouldn’t expect substantial differences in estimation performance.
Parameter Recovery

Average RMSE for item parameters within a replication for MH-RM (x-axis) and BA-EM (y-axis).

Takeaway: MH-RM and BA-EM are producing comparable results.
Ramsay Curve Recovery: Normal $g(\theta|\eta)$

- estimated RCs by replication shown in different colors
- true $g(\theta|\eta)$ shown in black
Ramsay Curve Recovery: Skewed $g(\theta|\eta)$

- estimated RCs by replication shown in different colors
- true $g(\theta|\eta)$ shown in black
Ramsay Curve Recovery: Bimodal $g(\theta|\eta)$

- estimated RCs by replication shown in different colors
- true $g(\theta|\eta)$ shown in black
Standard Error Estimation

For item parameters, average standard errors (x-axis) and Monte Carlo standard deviations (y-axis).

Note: For all three $g(\theta|\eta)$ shapes, the MC standard deviations are slightly larger.
**Standard Error Estimation**

Examining coverage probabilities (based on 100 replications):

| $g(\theta | \eta)$   | 68\% coverage |         | 95\% coverage |         |
|---------------------|---------------|---------|---------------|---------|
|                     | $a$           | $c$     | all           | $a$     | $c$     | all     |
| Normal              | 0.66          | 0.60    | 0.63          | 0.94    | 0.92    | 0.93    |
| Skewed              | 0.63          | 0.63    | 0.63          | 0.93    | 0.93    | 0.93    |
| Bimodal             | 0.67          | 0.64    | 0.66          | 0.95    | 0.93    | 0.94    |

- $a$ = slope
- $c$ = intercept
- “all” = slopes and intercepts
Example: National Comorbidity Survey

Background: The National Comorbidity Survey (NCS, 1994) was a nationwide household survey of the United States.

Ten Likert items (4-categories) measured current emotional distress.

Example Items:
- During the past 30 days how often did you...
  - worry too much about things?
  - feel exhausted for no good reason?
  - feel hopeless about the future?

Response options: never (0), rarely (1), sometimes (2), often (3).
Distress Scale Results: Ramsay Curve

Note: there appears to be a sizeable group of respondents at $\theta = -3$. 
Distress Scale Results: Slope Estimates and Standard Errors

<table>
<thead>
<tr>
<th>Item</th>
<th>Slope Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.99</td>
<td>.06</td>
</tr>
<tr>
<td>2</td>
<td>1.96</td>
<td>.07</td>
</tr>
<tr>
<td>3</td>
<td>2.08</td>
<td>.05</td>
</tr>
<tr>
<td>4</td>
<td>2.82</td>
<td>.07</td>
</tr>
<tr>
<td>5</td>
<td>2.50</td>
<td>.07</td>
</tr>
<tr>
<td>6</td>
<td>2.51</td>
<td>.06</td>
</tr>
<tr>
<td>7</td>
<td>2.86</td>
<td>.08</td>
</tr>
<tr>
<td>8</td>
<td>2.32</td>
<td>.08</td>
</tr>
<tr>
<td>9</td>
<td>2.35</td>
<td>.06</td>
</tr>
<tr>
<td>10</td>
<td>2.47</td>
<td>.07</td>
</tr>
</tbody>
</table>

Blue values correspond to RC estimation.
Black values correspond to IRT estimation with a normal distribution.
Outline

1 Introduction and Motivation
2 Review of Ramsay Curves
3 Review of the MH-RM Algorithm
4 Simulation Study
5 Empirical Application
6 Conclusion
Overall, results support the validity and utility of the MH-RM implementation.

Embedding RC-IRT in MH-RM (as opposed to EM) enables the computation of the observed information matrix. This, in turn, facilitates:

- test assembly
- limited-information goodness-of-fit testing
- differential item functioning
References I


Acknowledgements

This research is supported by grants from the Institute of Education Sciences (R305B080016 and R305D100039) and the National Institute on Drug Abuse (R01DA026943 and R01DA030466).

Thanks to Larry Thomas and Mark Hansen for providing datasets used in the empirical applications.

UCLA