Towards Bayesian Subject Model

Krzysztof Rusek

AGH University of Science and Technology

November 12, 2018
Simple subject model

\[ U_{ij} = \psi_j + \Delta_i + \nu_i X + \phi_j Y \]

\[ X, Y \sim N(0, 1) \]

Parameters

\[ \theta = (\psi_j, \Delta_i, \nu_i, \phi_j, \ldots) \]

Definition

Maximum likelihood estimator

\[ \hat{\theta} = \arg \max_\theta P(u|\theta) \]
1. \( U_{ij} \in \{1, 2, 3, 4, 5\} \sim N(\psi_j + \Delta_i, \sqrt{\nu_i^2 + \phi_j^2}) \)

2. Non unique solutions: how to partition variance among testers and PVSs?
1. $U_{ij} \sim Q(N(\psi_j + \Delta_i, \sqrt{\nu_i^2 + \phi_j^2})), Q()$ - a quantizer (e.g. ceil)

2. Use prior knowledge (expert, domain specific, etc) about parameters expressed as distribution.

**Example**

Bayesian model

1. $\nu \sim Gamm(\alpha_1, \beta_1)$
2. $\phi \sim Gamm(\alpha_2, \beta_2)$
Theorem (Bayes)

\[
P(\theta|u) = \frac{P(u, \theta)}{P(u)} = \frac{P(u|\theta)P(\theta)}{P(u)}
\]

Definition

Maximum a’posteriori estimator (MAP)

\[
\hat{\theta} = \arg \max_{\theta} P(\theta|u)
\]

MLE → MAP

\[
\log P(\theta|u) = \log P(u|\theta) + \log P(\theta) - \log P(u)
\]

Just add regularization given by \( \log \text{prior} \log P(\theta) \).
1. Based on TensorFlow, R like, GPU Accelerated.
2. Rich library of distributions (Normal, **QuantizedDistribution**).
4. Optimizers from TensorFlow for MAP, MLE ($\text{arg min}_\theta$).
5. Designed for Bayesian Inference (**MCMC**, tbd...).

**Example**

https://goo.gl/G4XR1C
Results
MLE - continuous

- True score vs. Recovered score
- True bias vs. Recovered bias
- True std vs. Recovered std
- True uppsilon vs. Recovered uppsilon

Krzysztof Rusek (AGH University of Technology)
Towards Bayesian Subject Model
November 12, 2018
MAP - continuous

- True score vs Recovered score
- True bias vs Recovered bias
- True std vs Recovered std
- True upsilon vs Recovered upsilon

Krzysztof Rusek (AGH University of Technology)  
Towards Bayesian Subject Model

November 12, 2018 10 / 1
MAP - quantized

- True score vs Recovered score
- True bias vs Recovered bias
- True std vs Recovered std
- True upsilon vs Recovered upsilon

Krzysztof Rusek (AGH University of Technology)
Towards Bayesian Subject Model
November 12, 2018
MAP as a simple extension of MLE
TensorFlow Probability allows fitting quantized distribution
Optimization poses numerical problems
Next step: MCMC for full posterior distribution (confidence intervals)