Computational Challenges in the Statistical Analysis of Stellar Evolution

Nathan Stein

with David van Dyk, Ted von Hippel, Steven DeGennaro, William H. Jefferys, Elizabeth Jeffery

October 20, 2009
The Model

- \( y_i = (y_{i1}, \ldots, y_{iJ}) \) = vector of magnitudes observed through \( J \) different filters
- \((M_{i1}, M_{i2})\) = primary and secondary mass of star \( i \)
- \( \theta \) = vector of cluster parameters
- \( G(M, \theta) \) = deterministic stellar evolution model
- Observational uncertainties \( \Sigma_i \) assumed known
- Gaussian errors:
  \[ y_i | M_i, \theta, \Sigma_i \sim N(\mu_i, \Sigma_i), \]
- For single-star systems, \( \mu_{ij} = G_j(M_{i1}, \theta) \)
- For main sequence-main sequence binaries,
  \[ \mu_{ij} = -2.5 \log_{10} \left( 10^{-G_j(M_{i1}, \theta)/2.5} + 10^{-G_j(M_{i2}, \theta)/2.5} \right) \]
- Mixture model to account for field star contamination
- Informative prior distributions on physical parameters
Posterior Correlations
Improving Mixing

- Widths of proposal distributions (Metropolis jumping rules) are automatically tuned during burn-in
- Parameters are transformed to remove linear correlations

\[ M_{i1} = U_i + \beta_{R,i}(R_i - \hat{R}_i) + \beta_{\text{age},i}(\theta_{\text{age}} - \hat{\theta}_{\text{age}}) + \beta_{[\text{Fe/H}],i}(\theta_{[\text{Fe/H}]} - \hat{\theta}_{[\text{Fe/H}]}) + \beta_{m-M_V,i}(\theta_{m-M_V} - \hat{\theta}_{m-M_V}) \]

\[ \theta_{A_V} = V + \gamma_{[\text{Fe/H}]}(\theta_{[\text{Fe/H}]} - \hat{\theta}_{[\text{Fe/H}]}) + \gamma_{m-M_V}(\theta_{m-M_V} - \hat{\theta}_{m-M_V}) \]
Improved Mixing

Before

After

\[ \text{iteration} \]

\[ \text{logAge} \]

\[ \text{[Fe/H]} \]
Removing Linear Correlations Is Not Enough

**Star 1**

- Primary mass: 1.004, 1.006, 1.008, 1.010
- Mass ratio: 0.0, 0.2, 0.4, 0.6

**Star 301**

- Primary mass: 0.900, 0.902, 0.904, 0.906
- Mass ratio: 0.0, 0.2, 0.4, 0.6

**Star 83**

- Primary mass: 1.02, 1.04, 1.06, 1.08, 1.10
- Mass ratio: 0.80, 0.90, 1.00

**Star 54**

- Primary mass: 1.125, 1.135, 1.145
- Mass ratio: 0.0, 0.2, 0.4, 0.6

---

**Star 1**

- Primary mass: 1.004, 1.006, 1.008, 1.010
- Mass ratio: 0.0, 0.2, 0.4, 0.6

**Star 301**

- Primary mass: 0.900, 0.902, 0.904, 0.906
- Mass ratio: 0.0, 0.2, 0.4, 0.6

**Star 83**

- Primary mass: 1.02, 1.04, 1.06, 1.08, 1.10
- Mass ratio: 0.80, 0.90, 1.00

**Star 54**

- Primary mass: 1.125, 1.135, 1.145
- Mass ratio: 0.0, 0.2, 0.4, 0.6
Power Law

Exponent: 1

Exponent: 2

Exponent: 4

Exponent: 6

Exponent: 8

Exponent: 10
More Correlations: 'Decorrelated' Masses?
Accelerating MCMC

- Want to sample $\pi(\theta)$ with $\theta = (\theta_1, \ldots, \theta_D) \in \Theta$
- Can obtain approximate sample (e.g., via trial run of inefficient MCMC sampler)
- Choose threshold $c \in (0, 1)$
- $r_{ij} =$ sample correlation of $\theta_i$ and $\theta_j$
- $I = \{i : |r_{ij}| \geq c$ for some $j \neq i\}$
- $M = |I|$
- $\theta = (\theta[I], \theta[-I])$
- $\{w_1, \ldots, w_M\}$ are linearly independent eigenvectors of $\text{cov}(\theta[I])$
- $\{w_i\}$ forms orthonormal basis for $M$-dimensional subspace of $\Theta$
- $W = M \times M$ matrix with columns $w_i$
- Alternative parameterization $\phi = W^T \theta[I]$
Accelerating MCMC: Algorithm

New MCMC scheme:

1. Update $\theta^{(t+0.5)} = \text{MCMC}(\theta^{(t)})$
2. Set $\phi^{(t+0.5)} = W^T \theta^{(t+0.5)}$
3. Draw $\phi^{(t+1)} \sim \pi(\phi | \theta^{(t+0.5)})$ (e.g., via Metropolis within Gibbs)
4. Set $\theta^{(t+1)} = W\phi^{(t+1)}$ and $\theta^{(t+1)}_{[-I]} = \theta^{(t+0.5)}_{[-I]}$
Accelerating MCMC: Illustration
Accelerating MCMC: Illustration
Accelerating MCMC: Illustration
Accelerating MCMC: Illustration
Accelerating MCMC: Illustration
Different Random Seeds

- logAge
- [Fe/H]
- modulus
- absorption
Multiple Modes
Multiple Modes

Primary Mass: Star 28

Mass Ratio: Star 28

Primary Mass: Star 35

Mass Ratio: Star 35
Multiple Modes