A Unified Theory of Photometric Redshifts

Tamás Budavári
The Johns Hopkins University
Haunting Questions

- Two classes of methods: empirical and template fitting
  - Why are they so different? Anything fundamental?

- Empirical techniques don’t work beyond the training set
  - Why not signal if query point is away from the training points?
  - Why can we only use the same observables for the estimation?
  - Why don’t they ever use the photometric errors?

- Template fitting has modest success
  - Why just use discrete sets and not continuous manifolds?
  - Why should we trust their estimates beyond our test set?

- People only estimate photometric redshifts
  - Why subsequent analyses to estimate other dependent properties?
Photometric Properties

- The “Photometric Inversion” problem
  - Constrain various properties consistently
  - Propagate their uncertainties and correlations

- Estimates are secondary
  - Should provide probability density functions instead
  - Scientific analyses should be using the full PDFs

- Make best use of observations
  - Next generation surveys are photometric only
(Yet another) History of life as we know it...

Homo apriorius
Homo pragmaticus
Homo frequentistus
Homo sapiens
Homo bayesianis

Courtesy of D. Madigan (Rutgers)
Basic Concepts

- Training and Query sets with different observables

\[
T : \{x_t, \xi_t\}_{t \in T} \\
Q : \{y_q\}_{q \in Q} \\
M : \theta
\]

- Model yields observables for given model parameter
  - Prediction via \( p(x, y|\theta, M) \) and has prior \( p(\theta|M) \)
  - Also folds in the photometric accuracy

- We are after \( p(\xi|y_q, M) \)
Mapping Observables

- The model provides the transformation rule

\[ p(x|y_q, M) = \int d\theta \ p(x|\theta, M) p(\theta|y_q, M) \]

with

\[ p(\theta|y_q, M) = \frac{p(\theta|M) p(y_q|\theta, M)}{p(y_q|M)} \]

- Think empirical conversion formulae but better
  - For example, from \textit{UJFN} to \textit{ugriz} with errors
Relation of Properties

- Usually just assume a function $\xi = \hat{\xi}(x)$
  - Wrong! We know there are degeneracies...

- There is a more general relation $p(\xi|x)$
  - Usual restriction is $p(\xi|x) = \delta(|\xi - \hat{\xi}(x)|)$
  - Correct estimation

$$p(\xi|x) = \frac{p(\xi,x)}{p(x)}$$

- Straightforward to do via KDE or Voronoi
  - I didn’t say cheap or easy ;-)}
Properties of Interest

- The final density of interest is

\[ p(\xi | y_q, M) = \int dx \ p(\xi | x) \ p(x | y_q, M) \]

Recommended Mantra: “It is not a function. It is not a func…”

- If the result is uni-modal (not likely), and you insist

\[ \bar{\xi}(y_q) = \int d\xi \ \xi \ p(\xi | y_q, M) \]
**Boundary Effects**

- Measure of reliability is the prob of $q$ making the cuts

\[ P(W|y_q, M) = \int dx \; P(W|x) \; p(x|y_q, M) \]

- Include observables used in the selection criteria
  - Cannot use just colors, if there was a magnitude cut
  - Cannot use just fluxes, if cut on morphology

*Ask me later if interested!*
Empirical method

- Normal distributions, same quantities: $\bar{x}(\theta) = \theta$ and $\bar{y}(\theta) = \theta$
- With simple prior, the mapping is analytic, e.g., for flat

\[ p(x_t | y_q, M) = \int d\theta \ N(x_t | \theta, C_t) N(\theta | y_q, C_q) \]

Template fitting

- Generate training set from SEDs
- If no errors $p(x | \theta, M) = \delta(|x - \bar{x}(\theta)|)$

\[ p(\xi | y_q, M) \propto \sum_{t \in T} \delta(|\xi - \xi_t|) N(y_q | \bar{y}(\theta_t), C_q) \]

cf. maximum likelihood
It works!
Red Galaxies
Blue Galaxies
Advanced Methods

- Mapping observables via templates
  - Arbitrary but complete basis on wavelength range
  - Can be non-physical because the physics is in the prior

- Empirical relation of properties
  - Your favorite fitting function or
  - Preferably the more generic relation

- Imagine empirical fitting with templates!
  - It’s easy if you can... ;-)

11/19/2008  Tamás Budavári (JHU)
Summary

- Unified theory from first principles
  - Aims to be as general as possible
- Solve the photometric inversion problem
  - Provide no estimates but probability densities
- Traditional methods in the classical limits
  - But already go beyond the usual techniques
- More advanced algorithms coming soon!

- Submitted to ApJ, preprint is at ArXiv/0811.2600