Notes on Photo-Z

Andrew Miller

January 20, 2015

Bayesian photometric redshift estimation, Benitez (2000)

Overview of Bayesian methods for photometric red-shift.

- Setup: given data $D = \{C, m_0\}$ (colors and magnitude of reference band), compute probability of red-shift (posterior distribution).
- Include information from a library of "templates"
  - Fully observed spectra (with spectroscopic red-shift)
  - Photometry data (with spectroscopic red-shift)
- galaxy templates, T (types/classes of galaxies, characterized by spectral density)
- red-shift prior $p(z|m_0)$ is a function of magnitude of signal

Benitez (2000) (cont'd)

- The fully Bayesian photo-z distribution averages over templates

$$p(z|C, m_0) = \sum_T p(z, T|C, m_0) \quad (1)$$

$$\propto \sum_T \underbrace{p(z, T|m_0)}_{prior} \underbrace{p(C|z, T)}_{likelihood} \quad (2)$$

$$p(z, T|m_0) = p(T|m_0)p(z|T, m_0) \quad (3)$$

- Can extend "templates" to parameterized spectral models characterized by S (which is what I want to do)

$$p(z|C, m_0) = \int dSp(z, S|C, m_0) \quad (4)$$

$$\propto \int dS \underbrace{p(z, S|m_0)}_{prior} \underbrace{p(C|z, S)}_{likelihood} \quad (5)$$

(e.g. S are PCA/NMF weights or spectra params)

Benitez (2000) (cont'd)

- Their likelihood (if I understand it correctly) looks like this

$$p(C|z,T) = p(f_u, f_g, \dots, f_z|z,T) \tag{6}$$

$$\propto \frac{1}{\sqrt{F_{TT}(z)}} \exp\left(-\frac{1}{2}\sum_\alpha \frac{(f_\alpha - af_{T_\alpha})^2}{\sigma^2_{f_\alpha}}\right) \tag{7}$$

  - $f_u, \dots, f_z$ are the observed fluxes ($\alpha \in \{ugriz\}$)
  - $f_{T,\alpha}(z)$ is the vector of fluxes of template T at red-shift z
  - a is a magnitude parameter
  - $F_{TT}(z)$ is a normalizing factor they define

Questions

- None of the papers I've read so far go into specific detail about going from spectra to fluxes (for instance with SDSS filter responses).
  - Can we go from object characteristics (spectra, luminosity, distance, redshift, size, etc) all the way to the generation of an SDSS image featuring just that object (and to the fluxes in the DR7QSO catalog)? Dimensional analysis (for my sake)?
  - How are the fluxes in the DR7QSO dataset computed?
  - How to deal with luminosity/distance and the SDSS fluxes? Can I just assume a spectral density with arbitrary units $f(\cdot)$ that integrates to one, and this defines the distribution over colors (like normalized SDSS fluxes)?
- How smooth are the spectra themselves (how much smoothing is done when deriving the PCA basis)? Model spectra of quasars and galaxies?
- Clustering vs. Factor Analysis?
- What kind of model of individual spectra makes sense? What are desirable model properties?

Budavari, 2001 (cont'd)

High level:

> *For a given redshift, the photometric observation gives constraints on the possible underlying SED, since we expect to get back the measured photometric values by redshifting the SED and convolving it with the filter response function. This constraint obviously depends on the photometric system and, also, the redshift of the object as the rest-frame spectrum is sampled at different wavelengths.*

s

- describes an algorithm for reconstruction a quasar spectrum template from photometric observations and spectroscopic redshifts. It seems sorta like a dynamic K-means/EM algorithm (they add spectral types as needed), and does a decent job reconstructing bumps where the emission lines are.

Budavari, 2001 (cont'd)

The algorithm is: start with a collection of initial spectral templates (I believe these are rest frame.) $\psi_1(\lambda), \ldots \psi_K(\lambda)$

1. Categorize all photometric observations in the training set into one of these K categories. Which is the most likely template to describe each observation?
2. Repair the estimated SEDs of each object (does this mean just de-redshift them?)
3. Replace each reference templates $\psi_k(\lambda)$ with the mean of each of the repaired templates of that class (discovered in step 1). This is like computing the new mean in k-means.
4. Check to see if you need to add a new template or remove an existing template based on some statistical criterion.

My proposed method is basically a fully probabilistic version of this.

Budavari, 2001 (cont'd)

Some takeaways from this paper:

- They land on four classes of quasars (Fig 7), each of which has a slightly different distribution of red shift (figure 6)
- Again, this paper doesn't really clarify to me an important aspect - how exactly is the χ^2 objective defined? How do you go from templates, T to fluxes?
- How does "clustering" compare to a decomposition method like PCA/NMF?

Brescia (2013) - Empirical Method

- use a multi-layer perceptron (four layers) regression setting on a combination of SDSS (from the DR7QSO dataset, I believe), UKIDSS, and WISE photometric datasets, comparing photo-z performance on the following intersections:
  1. SDSS: 1.1×105;
  2. SDSS $\cap$ GALEX: 4.5×104;
  3. SDSS $\cap$ UKIDSS: 3.1×104;
  4. SDSS $\cap$ GALEX $\cap$ UKIDSS: 1.5×104;
  5. SDSS $\cap$ GALEX $\cap$ UKIDSS *cap* WISE: 1.4×104.

  The largest dataset combined 43 features (mostly band fluxes and magnitudes). The authors mostly discuss the multi-layer model and their training technique, which is L-BFGS and various rounds of cross validation. The authors note their model's inability to generalize to regions of the space for which they don't have data (particularly large magnitudes or out of range z values).

Brescia (2013) - Empirical Method (cont'd)

The authors outline a bunch of statistics to compare between methods, including the bias, sample stdev, median of absolute value of two quantities

- $\Delta z = (z_{spec} - z_{phot})$ (residuals)
- $\Delta z_{norm} = (z_{spec} - z_{phot})/(1 + z_{spec})$ (normalized residuals)

And a bunch of percentages of outliers and such based on those statistics. One is "catastrophic outliers", defined as individual samples where $|z_{norm}| > 2\sigma(z_{norm})|$ - outside of two sample standard deviations.