

A METHOD FOR THE SPECTROSCOPIC INFERENCE OF FUNDAMENTAL STELLAR PARAMETERS

APRIL 30, 2014

IAN CZEKALA¹ ET AL.

1. INTRODUCTION

- Where our technique fits in the ecosystem of stellar methods.
- Applicable to any kind of spectrum (long-slit, echelle, infrared, flux-calibrated or not)
- Examples of fields where accurate and unbiased stellar parameters are crucial. Exoplanets. T Tauri stars.

2. METHOD

Typically, the stellar properties we most desire are effective temperature T_{eff} , radius or surface gravity $\log g$, and metallicity $[\text{Fe}/\text{H}]$. Several high-quality libraries of synthetic spectra now exist (PHOENIX, Kurucz, more being developed for GAIA) and are parameterized by these fundamental stellar parameters, which together, we call $\vec{\theta}_{\star, \text{grid}}$. Synthetic libraries provide high-resolution spectra which span a range of $\vec{\theta}_{\star, \text{grid}}$ covering the main sequence of spectral types. The library is typically specified on a grid of equal spacing in T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$.

In addition to these fundamental parameters, a star also has several observed properties that are a function of its kinematics, geometry, and location in our galaxy: projected rotational velocity $v \sin i$, line of sight velocity v_z , total extinction A_V , and solid angle Ω . We call these secondary parameters $\vec{\theta}_{\star, \text{post}}$, because we can model their effects on the synthetic spectra in a “post-processing” step by convolution with an appropriate kernel or multiplication with a smooth function. Together, we call $\vec{\theta}_{\star} = \{\vec{\theta}_{\star, \text{grid}}, \vec{\theta}_{\star, \text{post}}\}$.

iczekala@cfa.harvard.edu

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street MS 10, Cambridge, MA 02138

Using a synthetic stellar library as a back-end and applying post-processing techniques, we attempt to forward-model the observed stellar spectra in order to determine the best-fit $\vec{\theta}_{\star}$. Because properly sampled spectra have correlated pixels, we present a framework which self-consistently models these correlations.

Symbol	Description
i	index specifying a pixel
λ_i	wavelength corresponding to a given pixel i
$\vec{\theta}_{\star, \text{grid}}$	fundamental stellar parameters, T_{eff} , $\log(g)$, $[\text{Fe}/\text{H}]$, $[\alpha/\text{Fe}]$ that parameterize a synthetic spectrum from the grid
$\vec{\theta}_{\star, \text{post}}$	stellar parameters $v \sin i$, v_z , A_V , and R^2/d^2 that are applied during “post processing” of the synthetic spectrum
$\vec{\theta}_{\star}$	$\{\vec{\theta}_{\star, \text{grid}}, \vec{\theta}_{\star, \text{post}}\}$
$f_{\lambda, \text{inst}}(\lambda)$	data spectrum
$f_{\lambda, \text{synth}}(\lambda)$	synthetic spectrum
c_0, c_1, \dots	Chebyshev polynomial coefficients for residual flux calibration
$\vec{\theta}_N$	the set of nuisance parameters composed of $\{c_0, c_1, \dots, c_N\}$
$k(i \vec{\theta}_N)$	polynomial function that multiplies a synthetic spectrum
$\vec{\theta}$	the parameters $\{\vec{\theta}_{\star, \text{grid}}, \vec{\theta}_{\star, \text{post}}, \vec{\theta}_N\}$ that completely describe a model spectrum
D_i	data flux for a given pixel, $D(\lambda_i)$
\mathbf{D}	data vector comprised of all D_i , $i = \{0, \dots, N\}$
M_i	model flux for a given pixel, $M(\lambda_i \vec{\theta})$
\mathbf{M}	model vector comprised of all M_i , $i = \{0, \dots, N\}$
σ_i	Poisson noise for a given pixel i
R_i	residuals $D_i - M_i$
\mathbf{R}	residual vector $\mathbf{D} - \mathbf{M}$

Table 1
 Nomenclature used in this document

2.1. Interpolation from synthetic spectral libraries

Creating a new spectrum with a specific $\vec{\theta}_{\star, \text{grid}}$ requires either synthesizing a new spectrum using radiative transfer codes or interpolating from nearby grid points in the library. Because it is computationally expensive to synthesize a spectrum at high resolution over a wide wavelength range, we choose to interpo-

late. Spline interpolation for spectra (Husser 2012).

TODO: Implement band-limited interpolation from the synthetic grid using spline interpolation. Caution: As is, 100K/0.5 logg spacing in $\bar{\theta}_{*,\text{grid}}$ may not be Nyquist sampling.

TODO: Investigate Tom Loredó's mentions of "Cosmological Bayesian Emulator" for using Gaussian processes for interpolating from a simulation grid.

2.2. Post-processing synthetic spectra to match the data

Libraries of synthetic spectra are generally computed at high resolution ($R \geq 100,000$), sampled with many pixels per resolution element, and do not include any rotational broadening or account for any instrumental effects. In order to transform a raw synthetic spectrum interpolated from the grid into one that matches the spectrum of a real star, we must post-process the spectrum to account for these secondary effects. The projected rotational velocity of the star, parameterized by $v \sin i$, broadens spectral line profiles and is mathematically described by convolution with a kernel $g_{v \sin i}(\lambda)$. UV, optical, and infrared spectra are acquired using a spectrograph which disperses light onto a CCD, defining a specific resolution and sampling rate for the spectrum. Because the resolution of the spectrograph and spacing of the pixels are different from the synthetic spectral library, we must convolve the raw spectrum with the line spread function (LSF) of the instrument $g_{\text{LSF}}(\lambda)$ and resample to the exact pixels of the CCD. In wavelength space, the $v \sin i$ and LSF operations are represented by the convolution of the synthetic spectrum with these two kernels

$$f_{\lambda,\text{inst}}(\lambda) = g_{v \sin i}(\lambda) \otimes g_{\text{LSF}}(\lambda) \otimes f_{\lambda,\text{synth}}(\lambda) \quad (1)$$

Using the convolution theorem, we can rewrite these operations as multiplications in Fourier space

$$F_{\lambda,\text{inst}}(s) = G_{\text{LSF}}(s) G_{v \sin i}(s) F_{\lambda,\text{synth}}(s) \quad (2)$$

where $G \leftrightarrow g$ and $F \leftrightarrow f$ are Fourier transform pairs. In order to use the fast Fourier transform (FFT) to execute these operations using discrete samples of the synthetic spectrum $f_{\lambda,\text{synth}}$, we must first resample $f_{\lambda,\text{synth}}$

so that it is on a uniform grid. In the case of a spectrum, the natural uniform grid is one that is equally sampled in velocity space, such there is an equal velocity shift Δv between pixels. This results in a wavelength grid that is linearly sampled in the logarithm of wavelength. Therefore, the Fourier coordinate s is the number of cycles per sampling interval, having units of inverse velocity [s/km]. Next, we multiply the Fourier transform of the synthetic spectrum by the Fourier transforms of the rotational velocity and line spread function kernels. Lastly, we do an inverse FFT to transform the modified spectrum $F_{\lambda,\text{inst}}$ back to wavelength space, where it is sampled at the wavelength locations of the pixels in the detector ($f_{\lambda,\text{inst}}$). When resampling the spectrum, we are careful to ensure band-limited interpolation using splines in order to prevent introducing non-physical high frequency structure into the spectrum.

Finally we apply corrections for interstellar extinction and Doppler shift. Interstellar extinction attenuates the spectrum by an amount A_λ , which is wavelength-dependent. We use the (name here) dust-extinction model, which is parameterized by extinction in the V photometric band, A_V . We correct for a radial velocity difference v_z between the star and the earth by Doppler shifting the model spectrum.

TODO: Convolution in wavelength space is a possibility. Allows LSF to change with wavelength. Pixel width effects are not important as long as the LSF of the instrument is adequately sampled ($\gtrsim 3$ pixels across the LSF). If our spectrum were not properly sampled, we would need to additionally convolve with a boxcar the width of the pixels.

TODO: Doppler shifting the wavelengths could be implemented as a phase-shift in Fourier space, but it's pretty fast as-is.

Due to small imperfections in the flux-calibration of the data spectrum, there may be broad regions of the spectrum which are offset from the true continuum level by a slight amount. A traditional approach to address this problem is to normalize both the data and the spectrum to the continuum level before comparison. While this may work well for hotter stars with well-defined continua, for cooler stars with large molecular features, normaliza-

tion may introduce artificial features into the spectrum when the pseudo-continuum is incorrectly placed. Instead, we choose to multiply the model spectrum by a low-order Chebyshev polynomial that is normalized to unity. The coefficients of the polynomial, c_0, c_1, \dots , will be free-parameters in our model that we will solve for. If the behavior of the spectrograph is characterized by observations of spectrophotometric standards, then we can put priors on the coefficients that will prevent overfitting of the polynomials.

With more uncertainty, the polynomial multiplication can also be used as a substitute for flux-calibration since the sensitivity function of an instrument is usually a smooth function with wavelength. When spectra are not flux calibrated, there is greater uncertainty about the properties of the spectrograph and wider priors must be used on the Chebyshev coefficients. In this case, the polynomial might destroy broad-scale information present in the spectrum by fitting it out, something that would be avoided with more accurate priors provided by the flux calibration for the model.

2.3. Likelihood function

To evaluate which parameters of the model \mathbf{M} fit the data set \mathbf{D} best we use a standard multidimensional Gaussian likelihood function which allows for covariance between data points. If the residuals vector is

$$\mathbf{R}(\vec{\theta}_*) = \mathbf{D} - \mathbf{M}(\vec{\theta}_*) \quad (3)$$

with N data points, then the likelihood function is

$$p(\mathbf{D}|\vec{\theta}_*) = \frac{1}{\sqrt{(2\pi)^N \det(C)}} \exp\left(-\frac{1}{2}\mathbf{R}^T C^{-1} \mathbf{R}\right) \quad (4)$$

and the log-likelihood function is

$$\ln[p(\vec{\theta}_*)] = -\frac{1}{2}\mathbf{R}^T C^{-1} \mathbf{R} - \frac{1}{2} \ln \det C - \frac{N}{2} \ln 2\pi \quad (5)$$

C is the covariance matrix which describes the covariance between pixels in the spectrum. For independent noise with noise per pixel σ_i , this

matrix is

$$C = \begin{bmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & \sigma_1^2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix} \quad (6)$$

and Equation 5 reduces to the familiar χ^2 form of a sum over the square of the residuals, weighted by the inverse variance of each data point

$$\ln[p(\vec{\theta}_*)] \propto -\frac{1}{2}\chi^2 = -\frac{1}{2} \sum_i \frac{R_i^2}{\sigma_i^2} \quad (7)$$

However, residuals from a spectroscopic fit are often correlated due to the convolution of the stellar rotation and LSF kernels. Additionally, systematic errors in the synthetic spectra (such as incorrect line strengths) will result in regions of highly correlated residuals, which requires that we use a covariance matrix with off-diagonal terms σ_{ij} and Equation 5, which includes a matrix product instead of a sum over independent pixels.

In order properly account for these correlated residuals, we must understand the covariance structure and parameterize it in our model. We seek to understand two types of covariance structure: the large scale global covariance structure that results from the stellar rotation and LSF kernels, and a regional covariance structure that is localized to areas where spectral line mismatch results in small patches of highly correlated residuals.

TODO: Tom Loredo was saying that if we have a model that perfectly fits the data then our residuals will be uncorrelated, regardless of the PSF/LSF. But is this true? Shouldn't all noise be correlated on some scale by the LSF? Correspond by email to clarify and show an example.

To account for the global covariance structure, we use a Matern 3/2 kernel.

TODO: Explore difference between using 3/2 Matern kernel and squared exponential kernel.

$$k_{\nu=3/2}(r, a, l) = a \left(1 + \frac{\sqrt{3}r}{l}\right) \exp\left(-\frac{\sqrt{3}r}{l}\right) \quad (8)$$

tapered for compact support using a Hann

window with width $r_0 = 6l$ (Why? Because this looked good).

Specific line covariance. Non-stationary covariance kernel.

$$k_G(x, x'|a, \mu, \sigma) = \frac{a^2}{2\pi\sigma} \exp\left(-\frac{[(x-\mu)^2 + (x'-\mu)^2]}{2\sigma^2}\right) \quad (9)$$

$$k(x, x'|h, a, \mu, \sigma) = \exp\left(\frac{-(x-x')^2}{2h^2}\right) k_G(x, x'|a, \mu, \sigma) \quad (10)$$

h is a "bandwidth." If h is small, then there will be high-frequency structure. If h is large, then only low-frequency structure will remain.

These are all parameters that we can sample

in. Also, there would be a set of these parameters for each line in the model.

The benefit mainly comes from simply modelling these systematic residuals to begin with, since a model with covariance is far more likely than forcing the fit, and far more justified than masking regions which do not fit.

2.4. Applications

Learnt covariance structure can be used to correct the models.

REFERENCES

Husser, T.-O. 2012, 3D-Spectroscopy of Dense Stellar Populations (Universitätsverlag Göttingen)