The Application of Artificial Neural Networks to Stellar Classification

Coryn A.L. Jones

Institute of Astronomy, Madingley Road, Cambridge, United Kingdom
email: calj@ast.cam.ac.uk

Mike Irwin

Royal Greenwich Observatory, Madingley Road, Cambridge, United Kingdom

Ted von Hippel

WIYN Telescope, National Optical Astronomy Observatories, Tucson, AZ

Abstract. We are working on a project to automate the multi-parameter classification of stellar spectra using Principal Components Analysis (PCA) and Artificial Neural Networks. We present here the usefulness of PCA as a form of spectral data compression, and our results to date of classification on the MK system.

1. Introduction

Traditional methods of stellar classification are slow and arguably subjective. A robust, automated method is an essential requirement if stellar classification is to continue to be an important tool in the face of increasingly large data acquisition rates. We have been investigating the use of Artificial Neural Networks (ANNs) with a Principal Components Analysis (PCA) front-end compression as a means of quantifying stellar spectral classification.

2. Data Source

We have scanned and reduced some 100 IIaO plates taken as part of the Michigan Spectral Survey (Houk 1984). This has provided a set of some 15000 spectra down to $V \sim 11$ with a spectral coverage of $3800\mathring{A} - 5200\mathring{A}$ at a two pixel resolution of about $3\mathring{A}$. This corresponds to 820 flux bins per spectrum. Of these, we have selected 5000 high quality spectra across all spectral types which have 2D classifications from Houk & Smith-Moore (1988) (and earlier volumes). These classifications are used as the "targets" to train our neural networks.
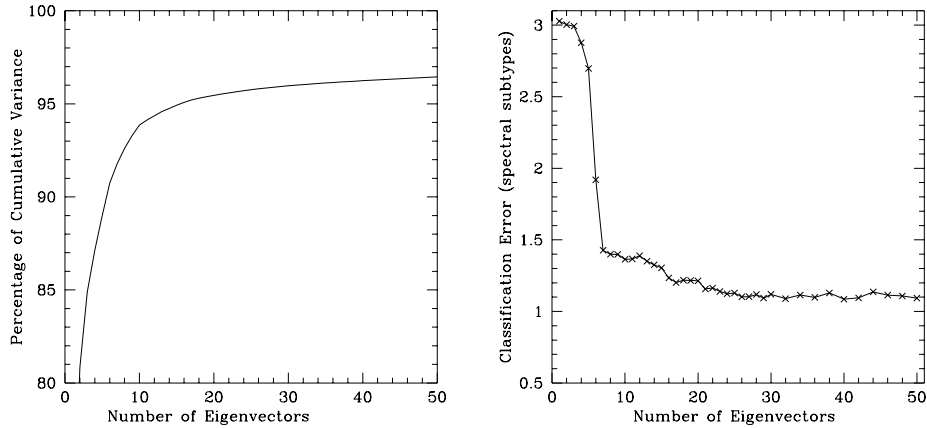
Figure 1. A: Variance of data explained by most significant eigen-
vectors. B: ANN error as a function of eigenvector representation of
spectra.

3. Artificial Neural Networks

Our principal classification tool is a neural network architecture known as a Back
Propagation Neural Network (see, e.g. von Hippel (1994); Jones (1996)). The
neural network consists of a number of weights which connect the inputs (the
spectrum), to the outputs (the classification parameters). Training the network
performs a non-linear mapping of the spectra onto the classification space and
sets the network weights. This process can be considered as a minimisation in
an N-dimensional space, where N is the number of resolution elements in the
spectrum. The network is trained using half of our data. The performance of
the network is evaluated by using it to classify the other half of the data set,
and comparing these classifications with the original catalogue classifications.

4. Principal Components Analysis

In terms of classification, a spectrum contains redundant information. Thus
we can increase the network training speed and over-determination factor by
compressing the data. It can be shown that a set of N-dimensional spectra
are optimally represented in a linear fashion by a set of basis vectors which
are the eigenvectors of the covariance matrix of the spectra (Storrie-Lombardi
et al. (1994)). Furthermore, the eigenvalues are proportional to the variance
explained by each eigenvector, so we can sort the eigenvectors into a significance
order. Figure 1a shows that just the first 50 eigenvectors reproduce over 96%
of the variance in the data. Whilst we do not use the eigenvectors directly to
classify the data (the spectral classification problem is not linearly separable),
it can be informative to analyse the components. Figure 2 shows the first 10
eigenvectors. Note that the continuum shape is spread over many components,
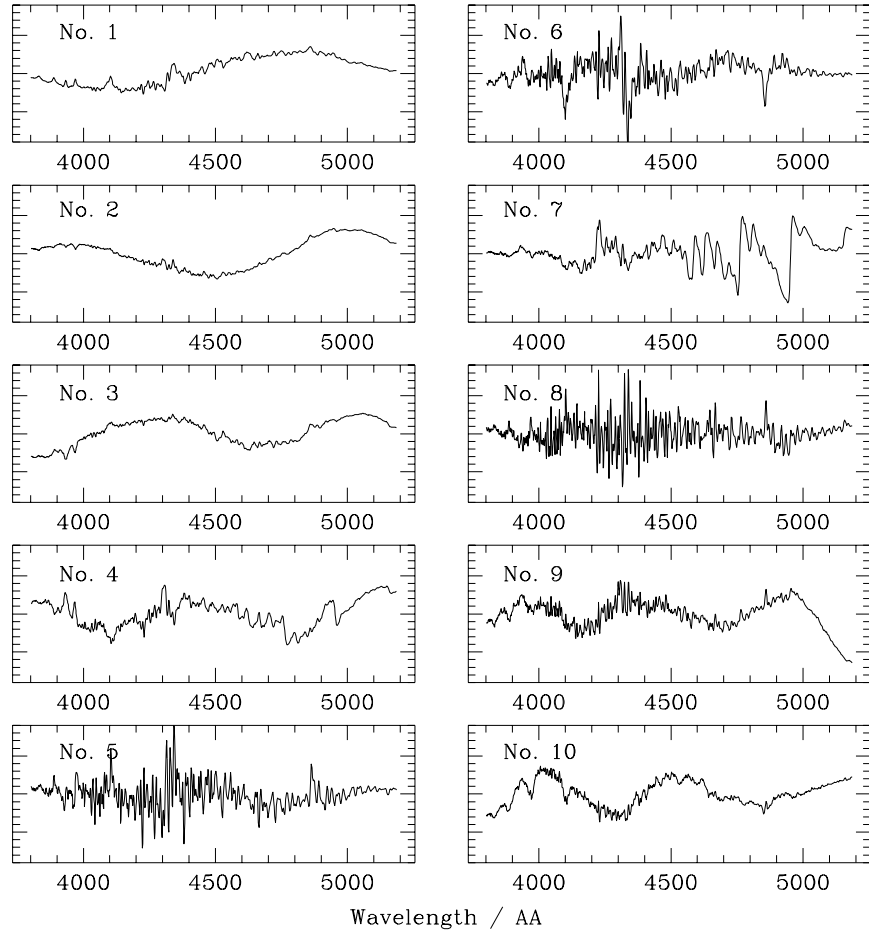as are the major lines of Ca H+K and the Hydrogen Balmer lines. Note also,

Figure 2.   First 10 normalised eigenvectors, plotted on the same scale. The sign of features is arbitrary.

that component 7 strongly represents the M stars on account of the strong TiO features.

5.   How Many Eigenvectors?

In order to investigate the effect of the number of eigenvectors used on the classification error, we trained a M:5:1 network to classify spectral types only. Figure 1b shows the classification error for a spectral representation using M eigenvectors, for M in the range 1-50. We see that using more than 25 eigenvectors is probably not necessary for the spectral type problem. Higher eigenvectors are presumably adding noise as fast as they are adding spectral information.
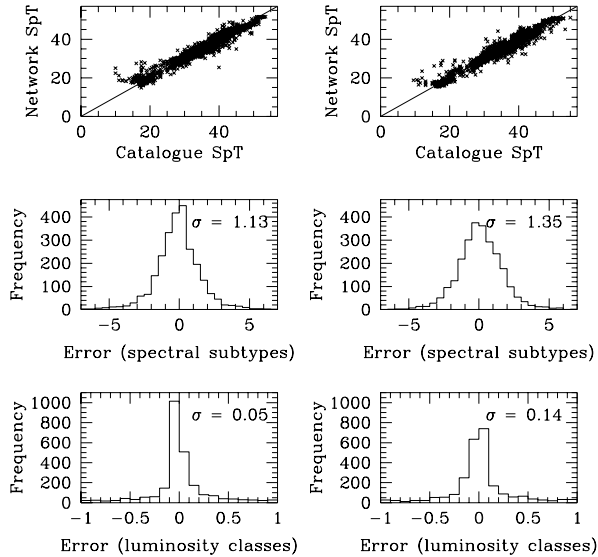
3

Figure 3.   ANN classification results. The left column is for line + continuum spectra and the right for line only. The spectral type (SpT) is coded onto a 1-57 range.

6.   Results

Using the first 50 principal components rather than the whole 820 flux bins as the ANN input yields a data compression factor of about 16. This number of components was used to develop a 50:5:2 network to simultaneously classify spectra into luminosity classes (III, IV or V) and the full range of spectral types. The results are shown in figure 3. We attempted classification with line only and line+continuum spectra. Despite potential inter-plate emulsion variations and reddening effects, the line+continuum classifications were of superior quality.

Acknowledgments.   CALJ would like to thank the ADASS Conference organisers for their generous support in enabling him to attend this conference.

References

Houk N. 1984, in Garrison R.F.. ed., *The MK Process and Stellar Classification*. David Dunlop Observatory, Toronto, p. 136.

Houk N. & Smith-Moore, 1988, *University of Michigan Catalogue of 2D Spectral Types for the HD Stars*, Vol. 4, and earlier volumes.

Jones C.A.L. et al. In preparation.

Storrie-Lombardi, M.C. et al., 1994, *Vistas in Astronomy*, **38**, 331.

von Hippel, T. et al., 1994, *MNRAS*, **269**, 97.