Data Analysis in Astronomy and Physics (SoSe22)

Instructor: PD Dr. Markus Röllig
TA: Dr. Christof Buchbender
TA: Craig Yanitski

Exercise Set 8

Due: **9:30 30 May 2022**
Discussion: **13:00 3 June 2022**
Online submission at via ILIAS in the directory Exercises / Übungen -> Submission of Exercises / Rückgabe des Übungsblätter

1. Error Calculations [40 points]

Suppose we are viewing a binary orbit face-on. The primary star has mass $2.19^{+0.43}_{-0.41} M_\odot$, luminosity $60.8^{+1.3}_{-1.2} L_\odot$, and effective temperature $6595^{+53}_{-58} K$. The secondary star has mass $1.62^{+0.26}_{-0.32} M_\odot$, luminosity $3.2^{+0.7}_{-0.8} L_\odot$, and effective temperature $4284^{+78}_{-73} K$. The distance to the binary is determined to be $5pc$. Compute the following properties including error.

hint: recall the solar values (without error)

$M_\odot = 1.989 \times 10^{30} kg$,

$m_\odot = -26.74$,

$L_\odot = 3.828 \times 10^{26} \frac{J}{s}$,

$R_\odot = 6.955 \times 10^{8} m$, and

$T_{eff,\odot} = 5780K$

a) What is the total mass of the binary? What is the reduced mass? **10 points**

hint: the reduced mass of M_1 and M_2 is $\mu = \frac{M_1 M_2}{M_1 + M_2}$

b) What is the radius of each star? **10 points**

hint: recall the equation for the luminosity of a star $L = 4\pi R^2 \sigma T_{eff}^4$, where σ is the Stefan-Boltzmann constant

c) What is the flux coming from each star? What is the total flux? What is the apparent magnitude of the binary system? **20 points**

hint: the flux is determined by $F = \sigma T_{eff}^4$, while apparant magnitude is given by $m = -2.5 log_{10}\left(\frac{F}{F_\odot}\right) + m_\odot$

2. PCA using covariance [60 points]

In this problem we will redo the Principal Component Analysis (PCA) as presented by Francis & Wills (1999) on a set of quasar data. The paper can be downloaded from this link. The data is available on the website as datafile: `quasar.dat`. The main result from the paper is shown in their Table 3:

Table 3. Results of Eigenanalysis – The Principal Components[a]

	PC1	PC2	PC3	PC4	PC5
Eigenvalue	6.4505	2.8157	1.5879	0.6257	0.5698
Proportion	0.496	0.217	0.122	0.048	0.044
Cumulative	0.496	0.713	0.835	0.883	0.927
Variable	PC1	PC2	PC3	PC4	PC5
log L_{1216}	0.053	0.535	−0.123	−0.029	−0.405
α_x	0.295	−0.198	0.079	0.485	−0.155
FWHM Hβ	−0.330	0.077	−0.357	−0.082	−0.141
Fe II/Hβ	0.341	−0.140	0.003	−0.487	−0.212
log EW [O III]	−0.310	0.016	0.255	0.394	−0.095
log FWHM C III]	−0.198	0.077	−0.623	0.054	0.402
log EW Lyα	−0.177	−0.502	−0.006	−0.143	0.033
log EW C IV	−0.336	−0.262	0.048	−0.050	−0.303
C IV/Lyα	−0.342	0.062	0.025	−0.074	−0.584
log EW C III]	−0.262	−0.413	−0.124	−0.176	−0.008
Si III]/C III]	0.342	−0.149	−0.018	−0.311	−0.116
N V/Lyα	0.231	−0.050	−0.573	0.107	−0.288
λ1400/Lyα	0.223	−0.351	−0.225	0.441	−0.216

We now carry through a PCA on the data of the quasar sample given in Francis & Wills (1999).

a) Read the paper! Remove data rows that have missing data. Create a table of the original data and compute mean value and standard deviation of each column. **10 Points**

b) Take the original data and put it into normalized or weighted form, so that the effect of different units is effectively removed. Normalize by the standard deviation! **10 Points**

c) Visually inspect the data after the normalization by plotting each column (x: data index, y: data value). Confirm (by eye) that each component is about normally distributed. **10 Points**

d) Construct the covariance matrix. This is a 13×13 symmetric matrix. **10 Points**

$$C_{ij} = \sigma_i \sigma_j$$

e) Compute the eigenvalues and eigenvectors of the covariance matrix. Plot the Eigenvalues against their number (index). Recreate Table 3 from Francis & Wills (1999). **10 Points**

f) Compute errors of the eigenvalues with a bootstrap analysis or jackknife. Use sample size of 10000. Plot the distributions for the first 5 eigenvalues. **10 Points**