

# 光谱处理

宋轶晗

LAMOST, NAOC

May 27, 2021



- 1 基本操作
- 2 光谱平滑
  - 为什么要平滑
  - 平滑的方法
  - 光谱插值
  - 光谱分辨率
- 3 谱线拟合
  - 认识谱线
  - 拟合谱线常用的函数
  - 谱线拟合的方法
  - 谱线拟合需要注意的地方
- 4 线指数计算
  - 什么是线指数
  - 指数例子
- 5 简单的并行处理
  - 并行的需求
  - 并行处理的一个简单的实现
  - 理想的并行处理



# Python 环境

- python 2 or python 3
- 系统自带 python 还是发行版本 (Anaconda) ?
- numpy, scipy, matplotlib, astropy

```
~/anaconda3/bin/pip install XXXXX
```

```
~/anaconda3/bin/python XXXXX
```

```
add command into ~/.bashrc or .bash_profile
```



## LAMOST 中分辨光谱结构

	No.	Name	Ver	Type	Cards	Dimensions	Format
1	0	Information	1	PrimaryHDU	91	()	
2	1	COADD_B	1	BinTableHDU	44	4175R x 5C	['E', 'E', 'E', 'I', 'I']
3	2	COADD_R	1	BinTableHDU	44	3710R x 5C	['E', 'E', 'E', 'I', 'I']
4	3	B-84117836	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
5	4	B-84117860	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
6	5	B-84117883	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
7	6	B-84117907	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
8	7	R-84117836	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
9	8	R-84117860	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
10	9	R-84117883	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']
11	10	R-84117907	1	BinTableHDU	42	4136R x 4C	['E', 'E', 'E', 'I']



# 读取 fits 文件

```
1 import numpy as np
2 import astropy.io.fits as fits
3
4 def load_data(filename):
5     with fits.open(filename) as hdulist:
6         hdu = hdulist[9] # block
7
8         flux = hdu.data.field('FLUX')
9         wave = hdu.data.field('LOGLAM')
10        wave = 10**wave
11    return wave, flux
```

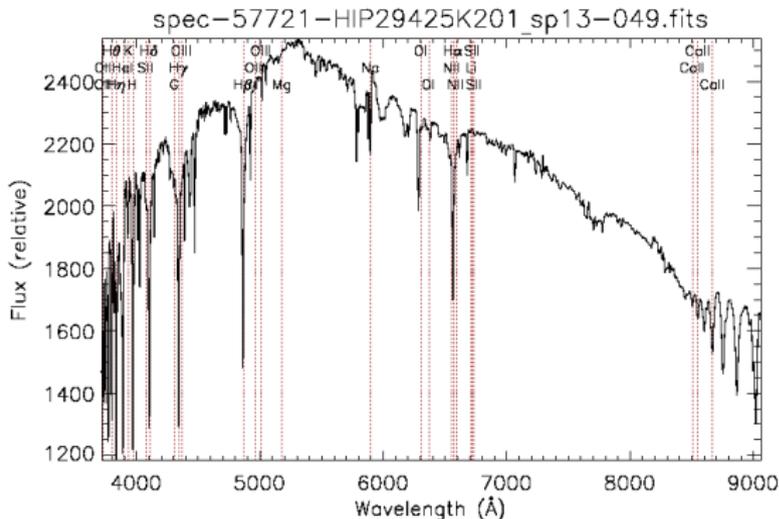


- 1 基本操作
- 2 光谱平滑
  - 为什么要平滑
  - 平滑的方法
  - 光谱插值
  - 光谱分辨率
- 3 谱线拟合
- 4 线指数计算
- 5 简单的并行处理



# 为什么要平滑

- 画图，好看
- 提高信噪比



STAR A0

$cz = 178.8 \pm -9999 \text{ km/s}$

RA = 95.55498, DEC = 23.56131

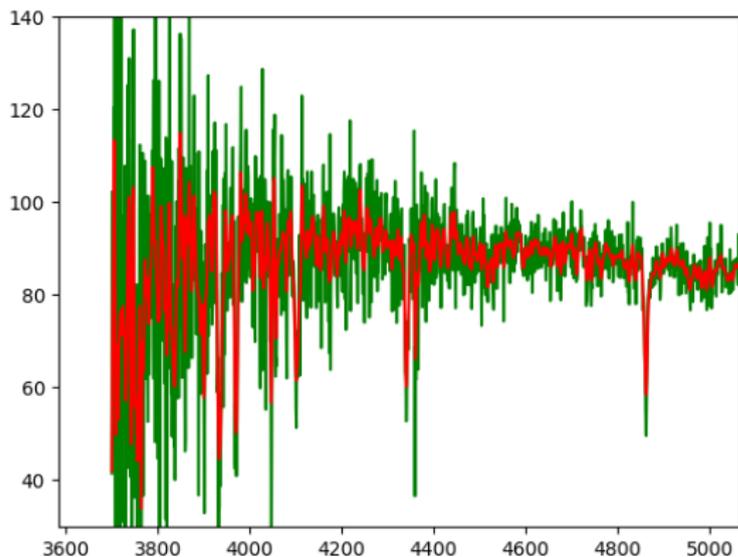
LAMOST DR8

Obs-Date: 20161128



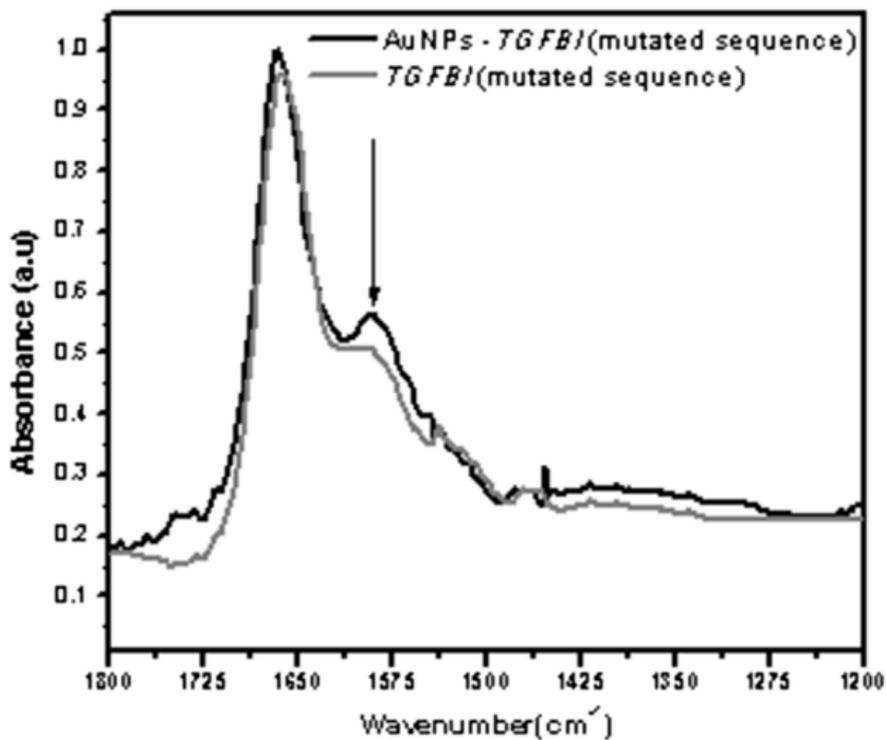
# 为什么要平滑

- 画图，好看
- 提高信噪比



# 为什么要平滑

- 画图, 好看
- 提高信噪比



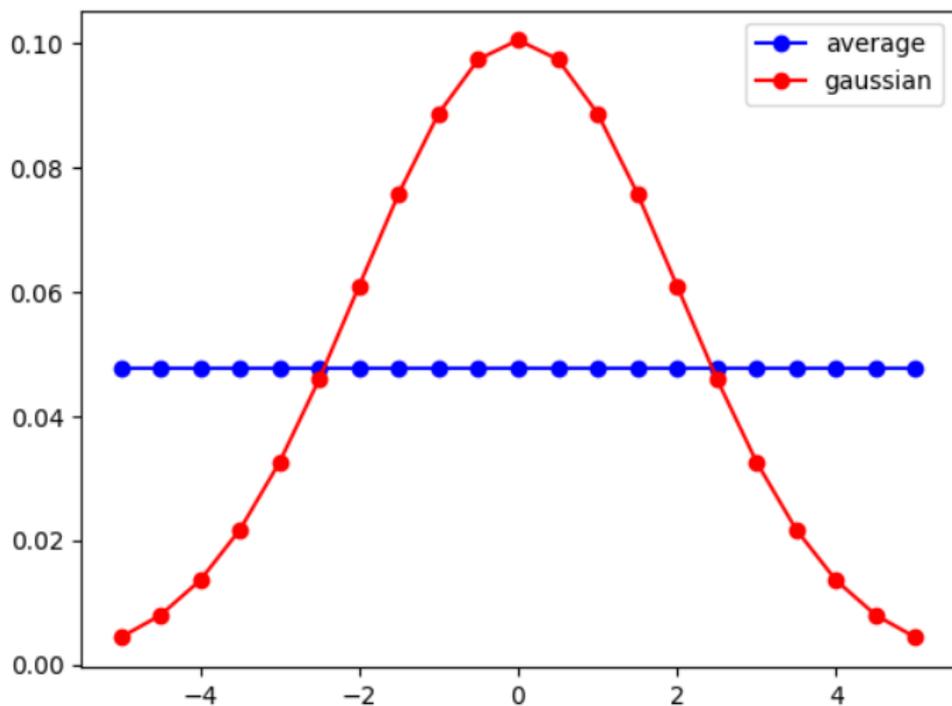
## 滑动窗口

$$S_i = \frac{1}{2m+1} \sum_{j=i-m}^{i+m} flux(j) \quad (1)$$

```
1 def moving_average(wave, flux, m=3):
2     n = len(flux)
3     s = np.zeros(n)
4
5     for i in range(len(flux)):
6         l = i-m
7         r = i+m+1
8         if l < 0:
9             l = 0
10        if r >= n:
11            r = n
12        data = flux[l:r]
13        s[i] = np.sum(data) / len(data)
14    return s
```



## 权重



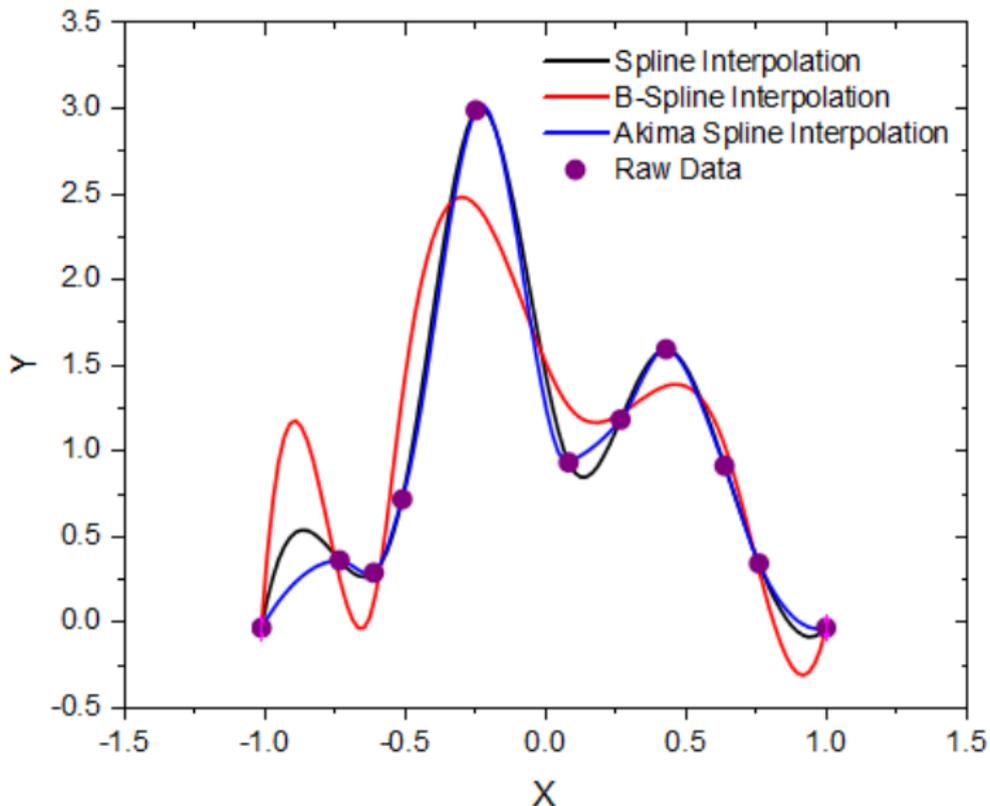
# 卷积

$$S = flux \otimes kernel \quad (2)$$

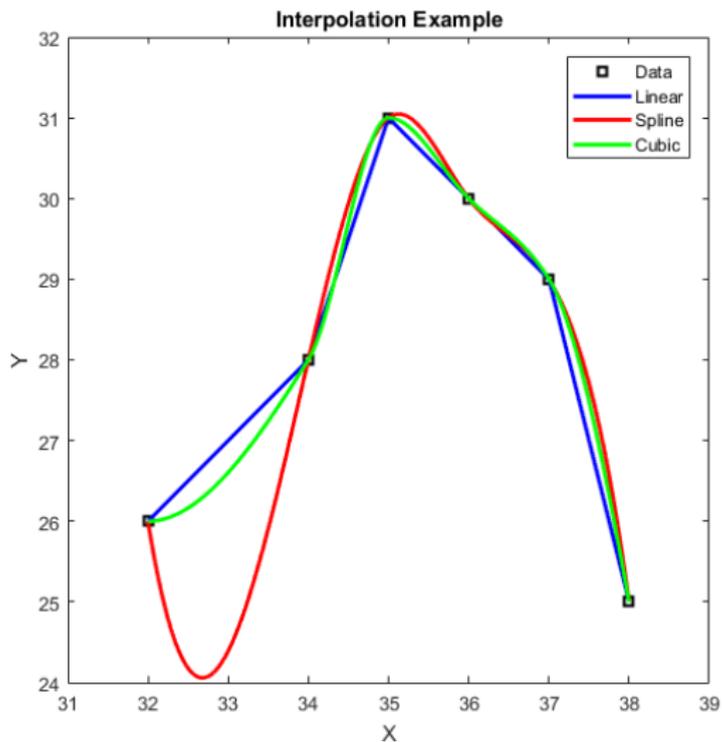
```
1 def conv(wave, flux, sigma = 2.0):  
2     x0 = np.median(wave)  
3     func = lambda x: np.exp(-0.5*(x-x0)**2/(sigma**2))  
4     kernel = func(x0)  
5     kernel = kernel / np.sum(kernel)  
6     y = np.convolve(flux, kernel, mode='same')  
7     return y
```



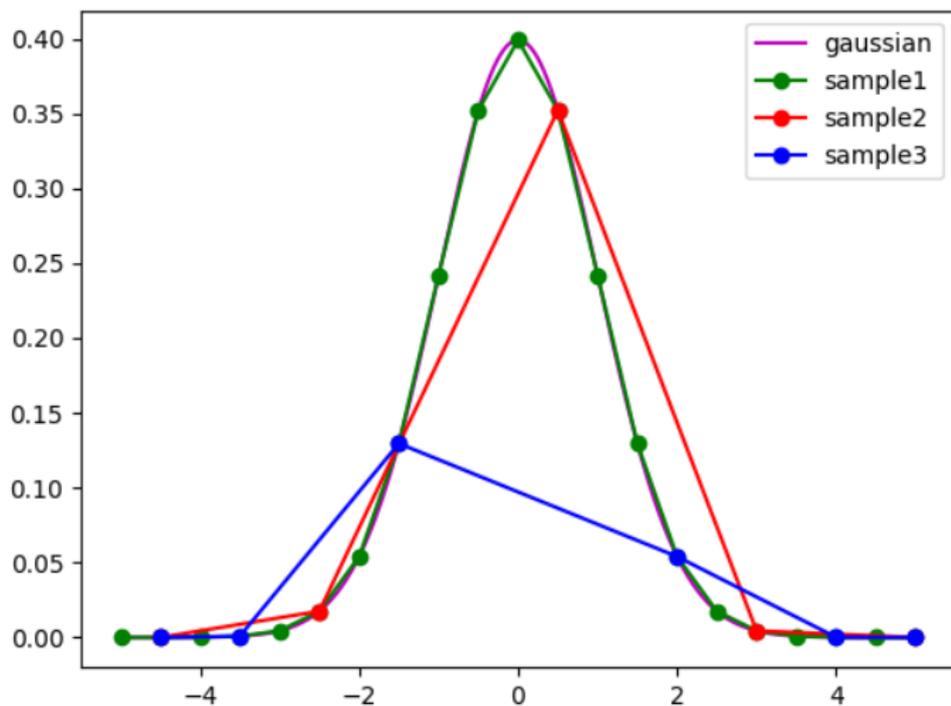
## 光谱插值



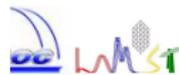
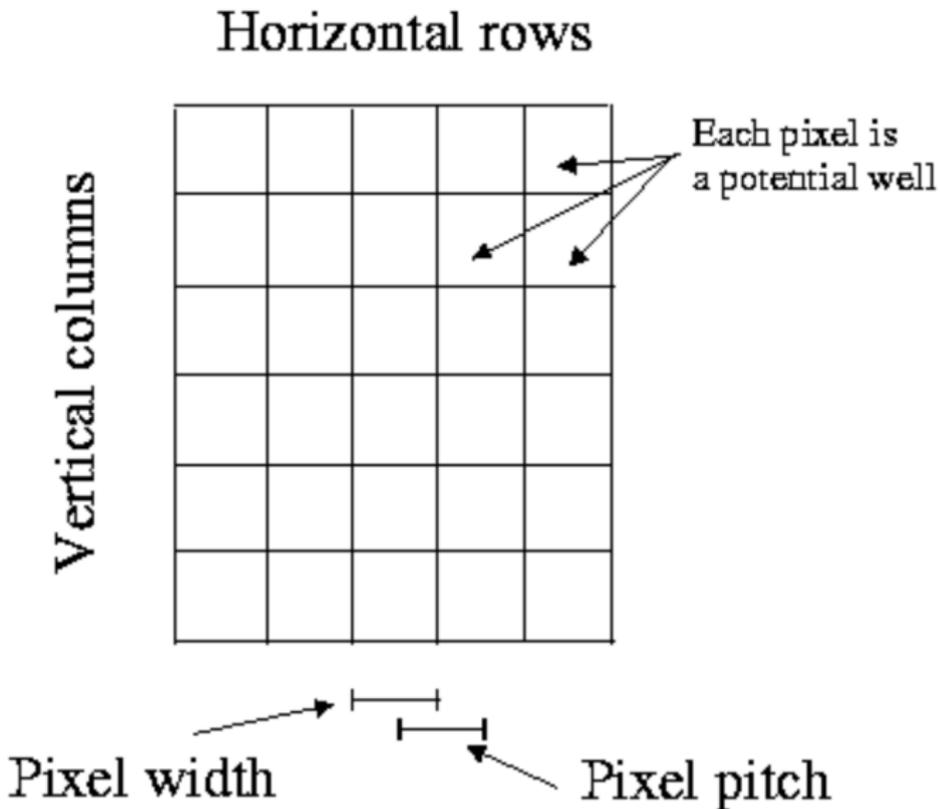
## 光谱插值



# 插值、采样



## 光谱分辨率

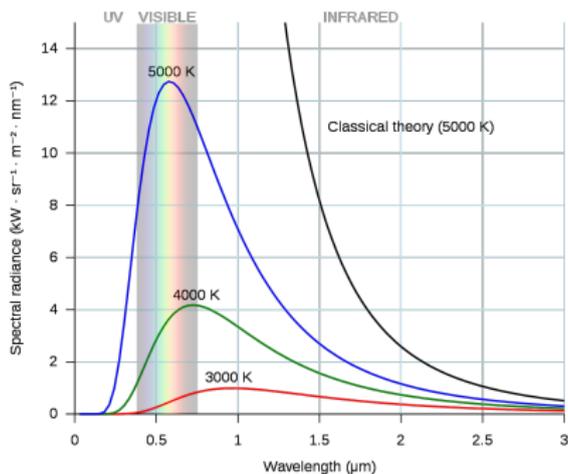


- 1 基本操作
- 2 光谱平滑
- 3 谱线拟合
  - 认识谱线
  - 拟合谱线常用的函数
  - 谱线拟合的方法
  - 谱线拟合需要注意的地方
- 4 线指数计算
- 5 简单的并行处理



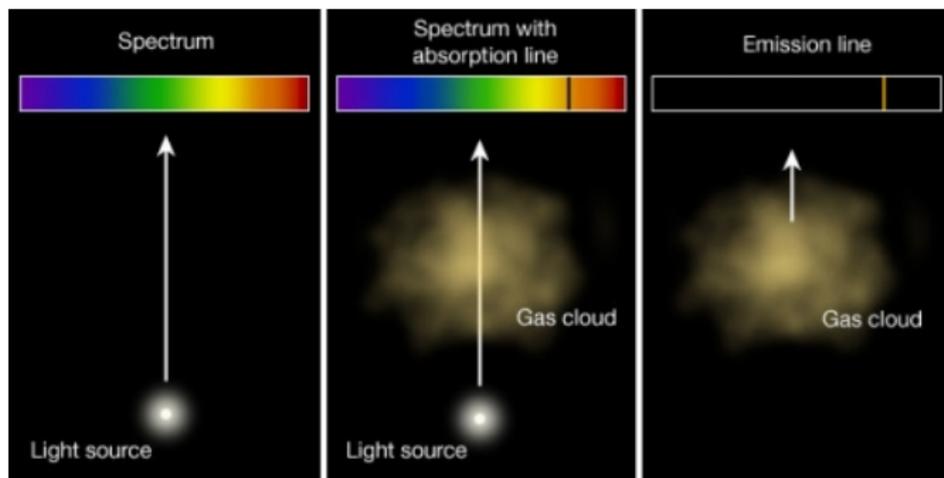
# 认识谱线

- 连续谱
- 发射线与吸收线
- 吸收线——能量被吸收比例



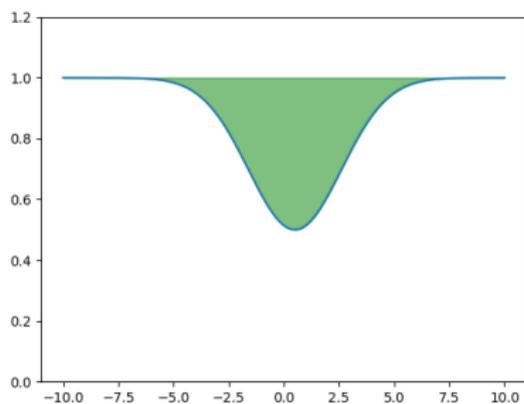
# 认识谱线

- 连续谱
- 发射线与吸收线
- 吸收线——能量被吸收比例



# 认识谱线

- 连续谱
- 发射线与吸收线
- 吸收线——能量被吸收比例



# 谱线形状

## 展宽

- 高斯型展宽：Doppler broadening, .....
- 洛伦兹型展宽：Life broadening, Pressure broadening, .....
- 仪器轮廓

## 最终

$$\text{line profile} = A \otimes B \otimes C$$



## Lorentzian

$$L(x; \gamma) = \frac{1}{\pi} \frac{\frac{1}{2}\gamma}{(\frac{1}{2}\gamma)^2 + x^2}, FWHM = \gamma \quad (3)$$

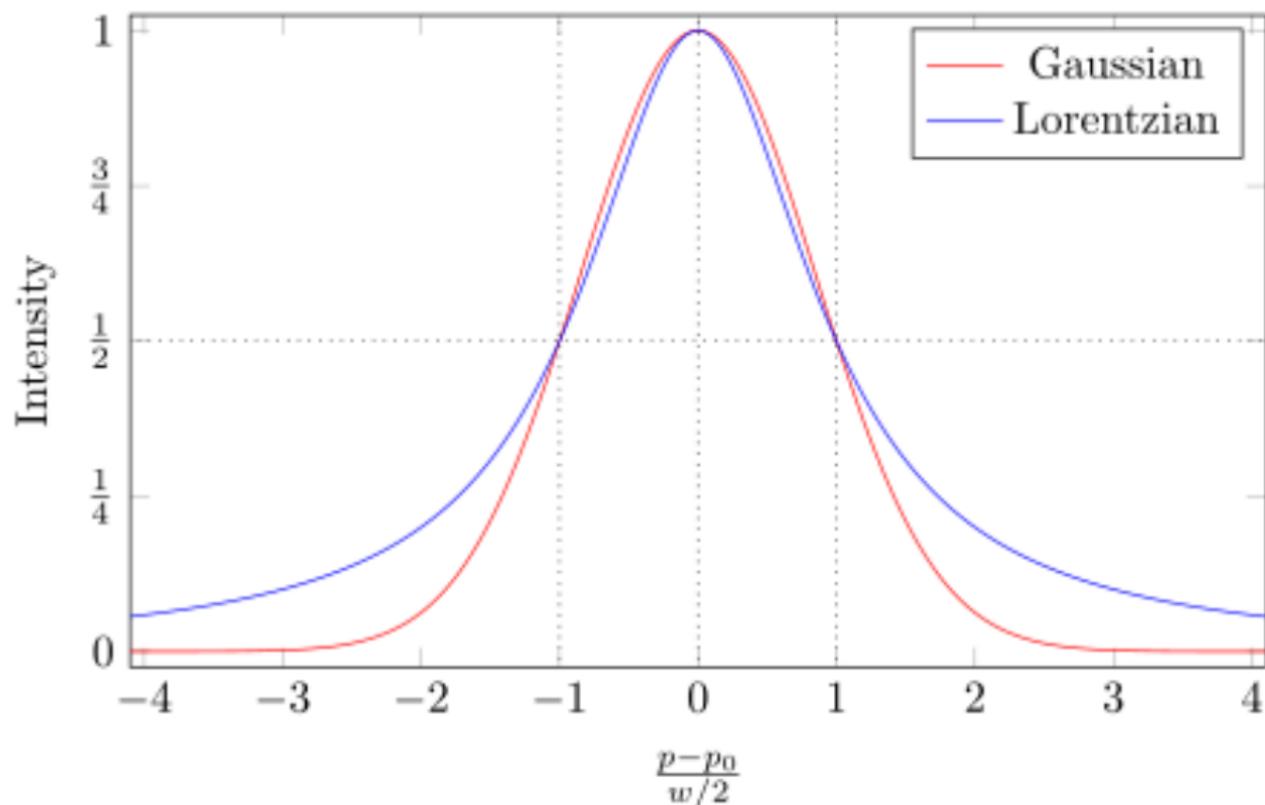
## Gaussian

$$G(x, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, FWHM = 2\sqrt{2\ln 2}\sigma \quad (4)$$

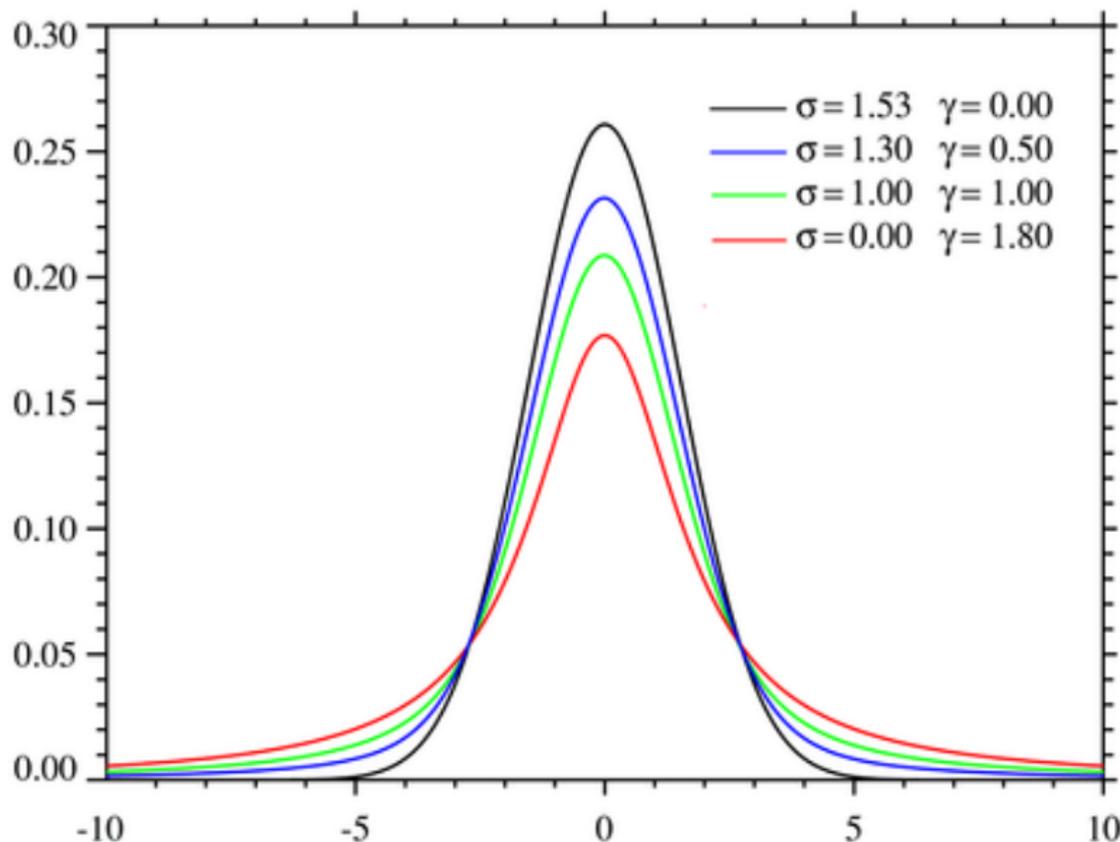
## Voigt

$$V(x; \sigma, \gamma) = \int_{-\infty}^{\infty} G(x'; \sigma) L(x - x'; \gamma) dx' \quad (5)$$

# Gaussian, Lorentzian



## Voigt

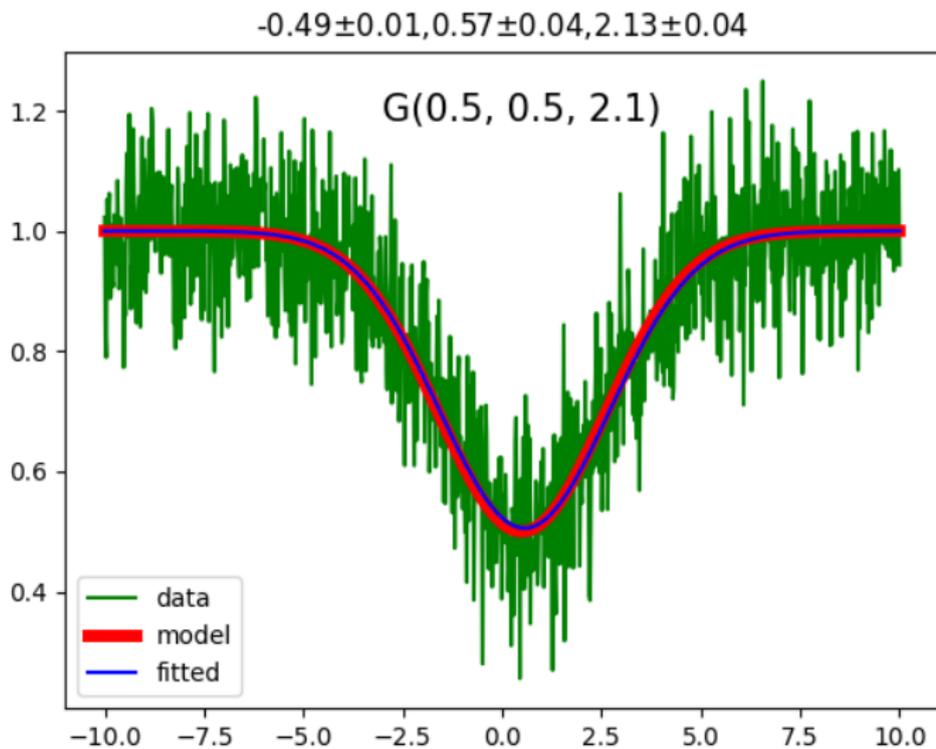


# 谱线拟合方法

```
1 import numpy as np
2 from scipy.optimize import curve_fit
3
4 def gauss_func(x, a, b, c):
5     return a*np.exp(-0.5*(x-b)**2/(c**2)) + 1
6
7 def fit_data(x,y,p0):
8     #p0 = [1,0.0,2.0]
9     popt,pcov = curve_fit(gauss_func, xdata=x,
10                          ydata=y,p0=p0)
11     err = np.sqrt(np.diag(pcov))
12
13     return popt, err
```



## 例子



# 谱线拟合需要注意的地方

- 线的轮廓决定选择的函数
- 初始参数非常重要
- 对拟合的参数进行合理性判断

## 说明

如果选错了拟合的函数，可能得到不稳定，或者错误的结果。



# 谱线拟合需要注意的地方

- 线的轮廓决定选择的函数
- 初始参数非常重要
- 对拟合的参数进行合理性判断

## 说明

拟合过程通过最小化实现，如果初始值设置不好，拟合结果会到另外一个极值点上。



# 谱线拟合需要注意的地方

- 线的轮廓决定选择的函数
- 初始参数非常重要
- 对拟合的参数进行合理性判断

## 说明

有的时候拟合会出现意想不到的结果，通过对拟合的参数合理的判断，来确定拟合的结果是否为我们想要的结果。



- 1 基本操作
- 2 光谱平滑
- 3 谱线拟合
- 4 线指数计算
  - 什么是线指数
  - 指数例子
- 5 简单的并行处理



# 什么是线指数

## 指数是什么

- 一个 数
- 一个特征
- 随便定义,  $A-B$ ,  $A/B$ ,  $A/B+C$
- 横向比较

## 优缺点

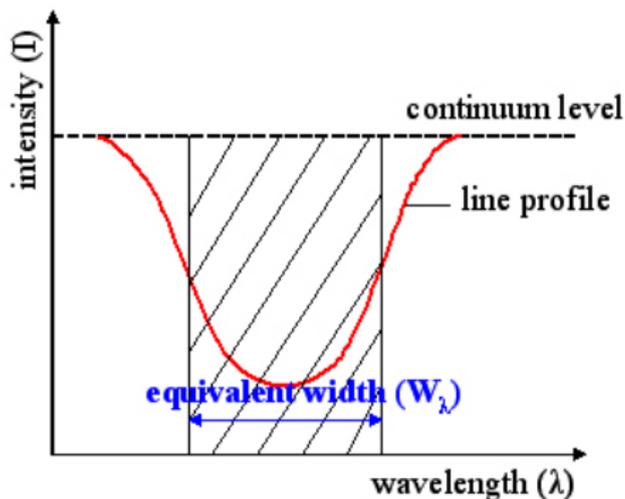
- 计算速度快
- 对分辨率等敏感
- 对信噪比不是非常敏感



# 等值宽度

wiki 定义

$$W_\lambda = \int 1 - \frac{F_\lambda}{F_c} d\lambda \quad (6)$$



## Lick 指数

## Lick-Index Calibration of GMOS

TABLE A1  
LICK INDEX PASSBAND DEFINITIONS

Index	feature passband		blue continuum		red continuum		units
$H\delta_A$	4083.500	4122.250	4041.600	4079.750	4128.500	4161.000	Å
$H\delta_F$	4091.000	4112.250	4057.250	4088.500	4114.750	4137.250	Å
$CN_1$	4143.375	4178.375	4081.375	4118.875	4245.375	4285.375	mag
$CN_2$	4143.375	4178.375	4085.125	4097.625	4245.375	4285.375	mag
$Ca4227$	4223.500	4236.000	4212.250	4221.000	4242.250	4252.250	Å
$G4300$	4282.625	4317.625	4267.625	4283.875	4320.125	4336.375	Å
$H\gamma_A$	4319.750	4363.500	4283.500	4319.750	4367.250	4419.750	Å
$H\gamma_F$	4331.250	4352.250	4283.500	4319.750	4354.750	4384.750	Å
$Fe4383$	4370.375	4421.625	4360.375	4371.625	4444.125	4456.625	Å
$Ca4455$	4453.375	4475.875	4447.125	4455.875	4478.375	4493.375	Å
$Fe4531$	4515.500	4560.500	4505.500	4515.500	4561.750	4580.500	Å
$Fe4668$	4635.250	4721.500	4612.750	4631.500	4744.000	4757.750	Å
$H\beta$	4847.875	4876.625	4827.875	4847.875	4876.625	4891.625	Å
$Fe5015$	4977.750	5054.000	4946.500	4977.750	5054.000	5065.250	Å
$Mg_1$	5069.125	5134.125	4895.125	4957.625	5301.125	5366.125	mag
$Mg_2$	5154.125	5196.625	4895.125	4957.625	5301.125	5366.125	mag



# S 指数

$$S = \alpha \frac{N_H + N_K}{N_R + N_V} \quad (7)$$

S-Index: The flux ratio of two bandpasses centered on the H (3968.5 Angstrom) and K (3933.7 Angstrom) Ca II emission cores and two continuum regions on either side. The S-Index includes contributions from both the photosphere and the chromosphere of the star. The Exoplanet Archive uses the Mt. Wilson S-Index.



- 1 基本操作
- 2 光谱平滑
- 3 谱线拟合
- 4 线指数计算
- 5 简单的并行处理
  - 并行的需求
  - 并行处理的一个简单的实现
  - 理想的并行处理



# 并行的需求

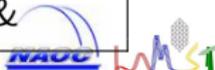
- 数据数量爆炸式增加
- 尽量利用机器资源
- 配置比较复杂的并不适合我们



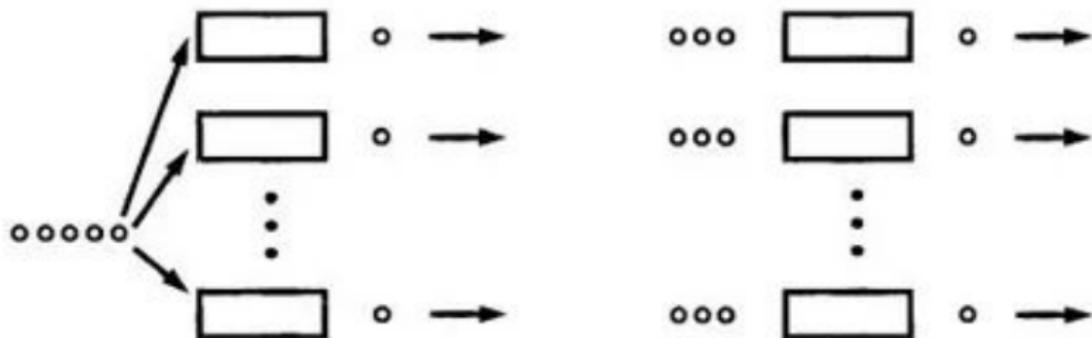
# 并行处理的一个简单的实现

```
1 import sys
2
3 args = sys.argv
4 base,rem = int(args[1]), int(args[2])
5
6 # a list of all data
7 candidates = load_all_candidate()
8 for ind,data in enumerate(candidates):
9     if ind % base != rem:
10         continue
11     work_to_do(data)
```

```
1 nohup python XXXX.py 3 0 1>out.0.1 2>out.0.2 &
2 nohup python XXXX.py 3 1 1>out.1.1 2>out.1.2 &
3 nohup python XXXX.py 3 2 1>out.2.1 2>out.2.2 &
```



# 理想的并行处理



# Question?

Thank you  
yhsong@nao.cas.cn



# 练习

## 练习 1

blue center	blue width	line center	line width	red center	red width	name
6725.0	50.0	6562.8	24.0	6425.0	50.0	Halpha

根据上面线指数定义，计算线指数，并比较光谱平滑之后的线指数。  
左右两段连续谱的流量与波长的中值连线作为局部连续谱。

## 练习 2

- 理解初始值，尝试初始值在什么区间能够得到正确的拟合结果。
- 谱线拟合，计算每个谱线的参数。

## 练习 3

并行处理上面两个练习。