P_Spline Performance

1) Runtime

The new **p_spline**.py with penalized normal equation has been run in Taurus machine with the following results.

Note: The time shown below excludes the time taken to perform loading of the data.

# pixels	Runtime (Taurus)
1 pixel	0.1469 seconds
307200 pixels	37.5675 seconds

2) P_Spline validation

a) Single pixel performance validation

 # pixel = 1

λ	# Knots	GCV	AIC	MSE
1	200	3.43E-06	2148.845922	0.003043836
0.1	200	3.42E-06	2243.061721	0.00274886
0.01	200	3.58E-06	2327.457311	0.002622919
0.001	200	3.69E-06	2356.900539	0.002609695
0.0001	200	3.71E-06	2361.409575	0.002609279

$$\mathrm{GCV}(\lambda) = \sum_{i=1}^{m} \frac{(y_i - \hat{y}_i)^2}{(m - \sum_{i=1}^{m} h_{ii})^2}.$$

$$\mathrm{AIC}(\lambda) = \mathrm{dev}(y;\, a, \lambda) + 2 * \dim(a, \lambda),$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$$

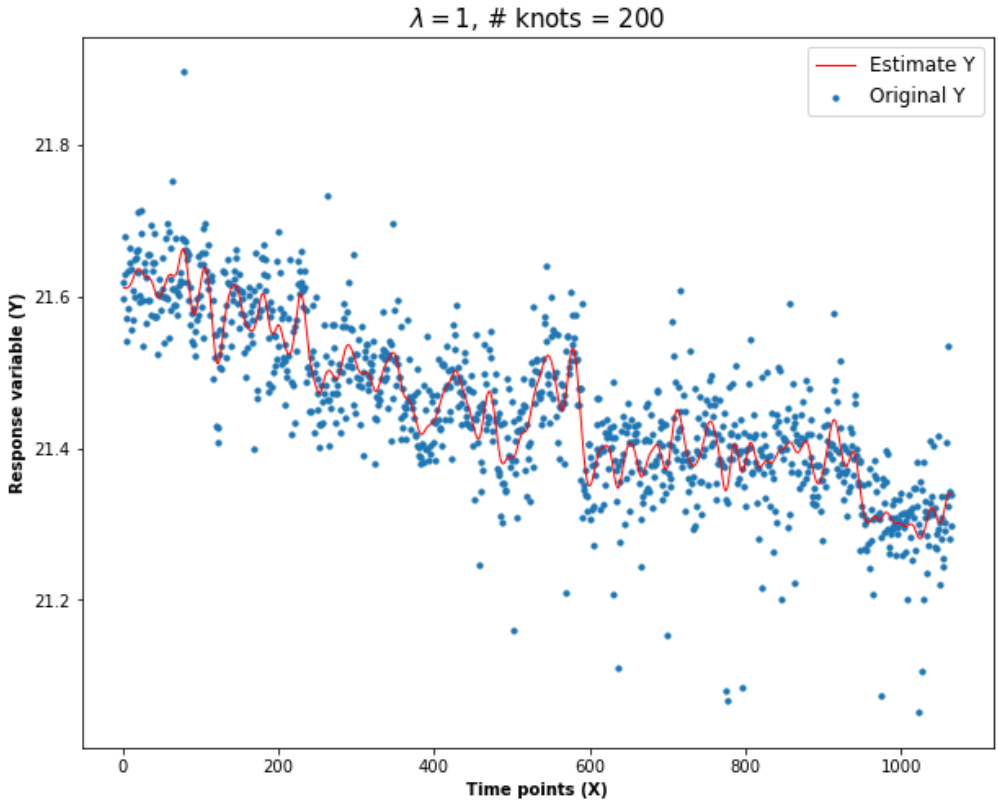
λ =smoothing parameter, **m**= number of observations

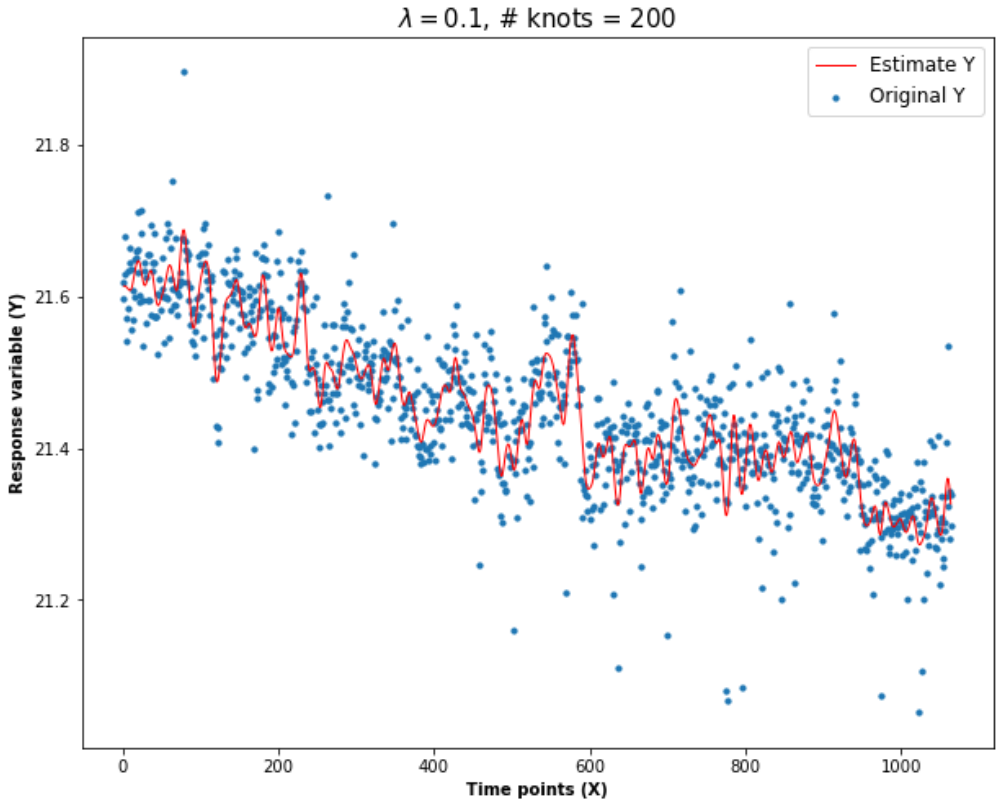
y = original values, **y_hat** = estimate of y

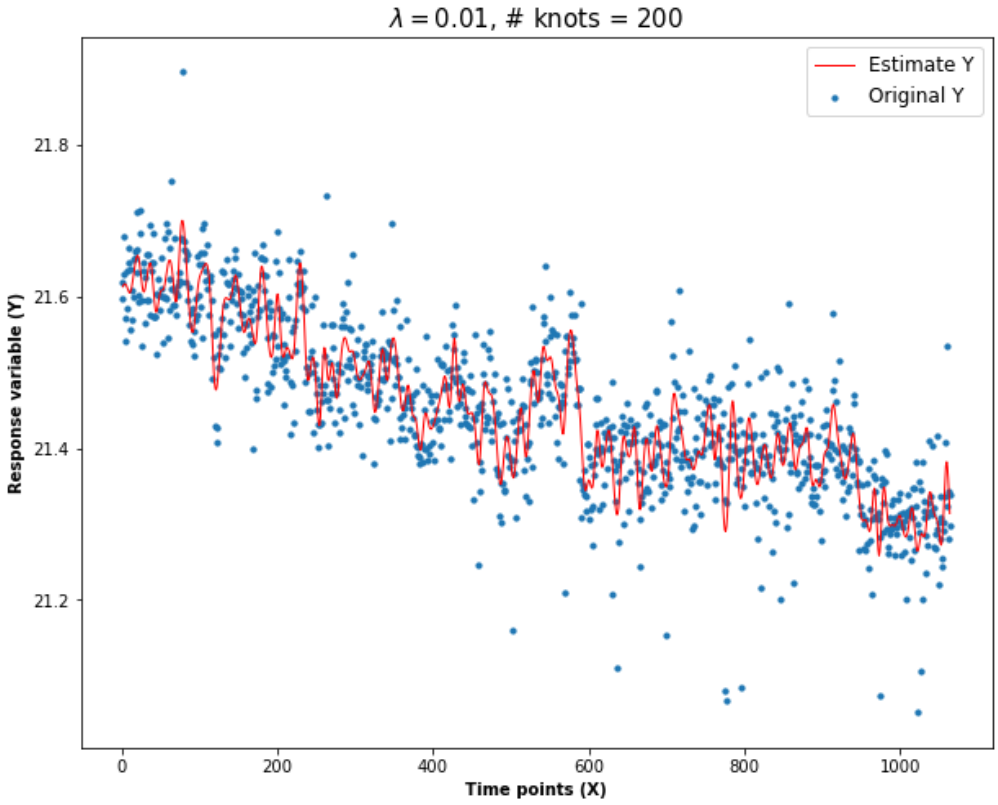
h= diagonal elements of the hat matrix H, **dim** = dimension of the vector of coefficients, a

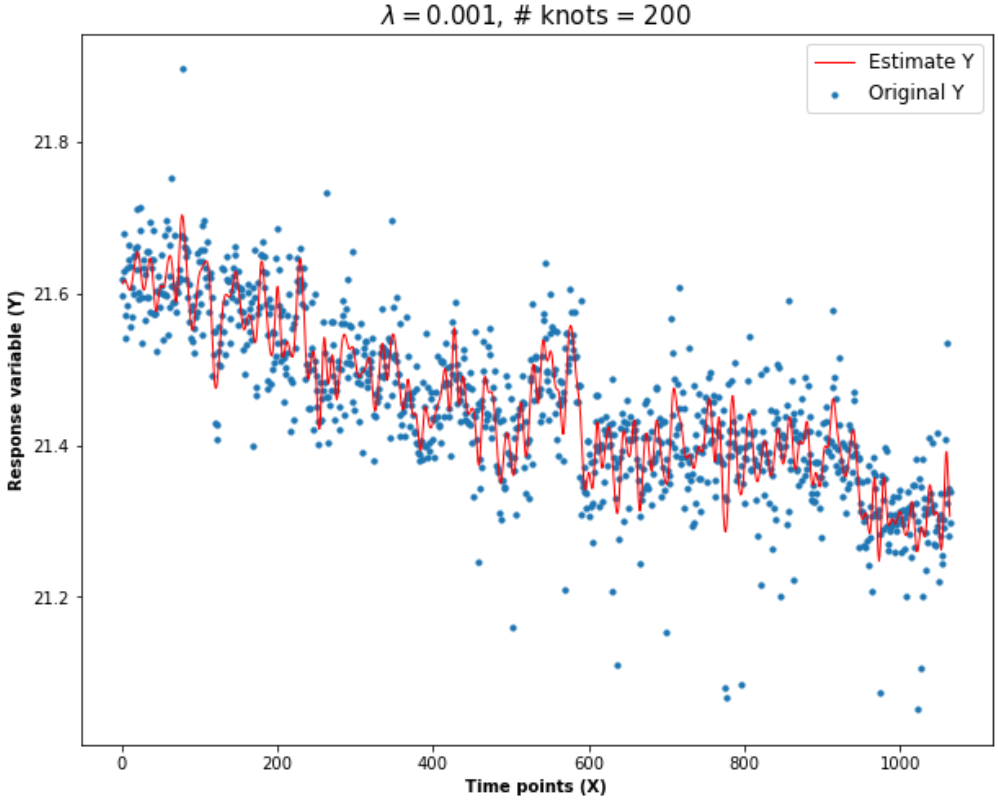
dev = deviance

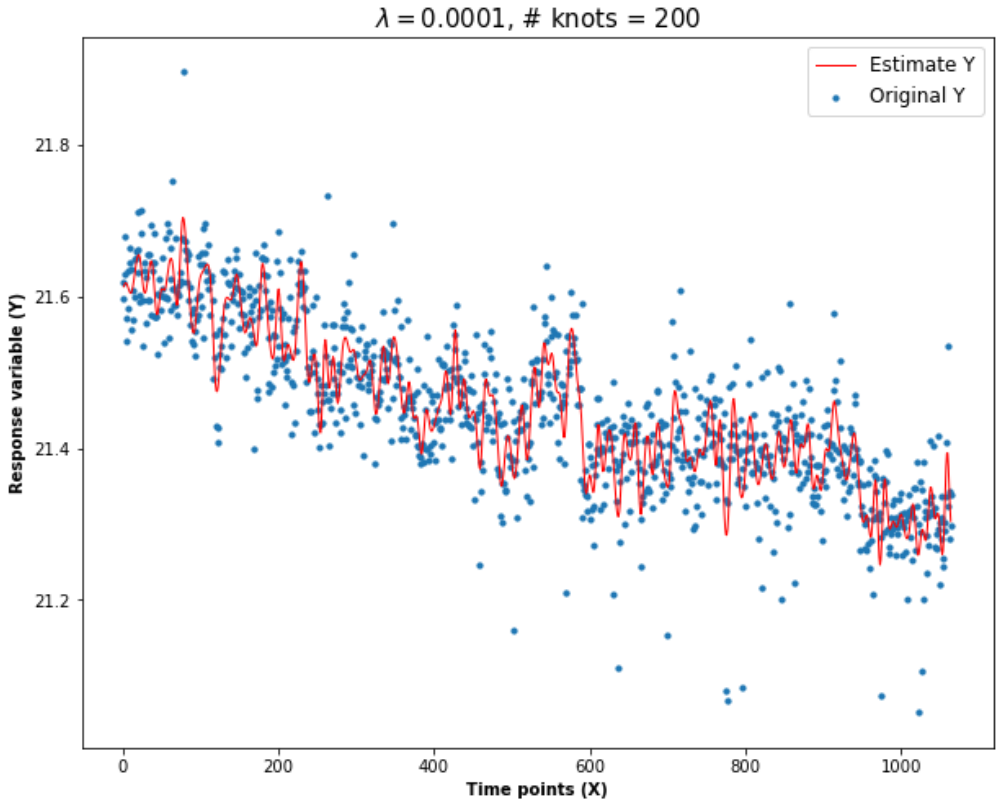
b) Single pixel figures






$\lambda = 0.01$, # knots = 200
Estimate Y
Original Y
Response variable (Y)
Time points (X)
21.8
21.6
21.4
21.2
0
200
400
600
800
1000


$\lambda = 0.001$, # knots = 200
Estimate Y
Original Y
Response variable (Y)
Time points (X)
21.8
21.6
21.4
21.2
0
200
400
600
800
1000



c) 307200 pixels performance validation

 # pixels = 307200

λ	# Knots	GCV	AIC	MSE
1	200	1.226327767	603019840.2	0.003543761
0.1	200	1.124027553	602822274.3	0.002940172
0.01	200	1.109265834	602724912.9	0.002642605
0.001	200	1.131362096	602712989.8	0.002606106
0.0001	200	1.136697776	602712573	0.002604819