2018-09-01

Polynomial Fitting

R. Steven Turley
Brigham Young University, turley@byu.edu

Polynomial Fitting

Version 1.0

R. Steven Turley

September 1, 2018

Contents

1 Introduction

Fitting a polynomial to data is a special case of linear least squares. Since I discussed that in detail in my Linear Least Squares article[1], so I won't repeat that information or those references here except to establish notation. Our goal is to find the coefficients b_i in the function

$$f(x;\vec{b}) = sum_{k=1}^{p} b_k x^{k-1} \tag{1}$$

that minimize the sum of the squares of the differences between the function f and the measured data.

2 Special Polynomial Formulas

In this section I will specialize the formulas derived in the Linear Least Squares article[1] for polynomials.

2.1 Finding Fit Parameters

In order to determine the fit parameters, we need to fill the matrix A and vector $\vec{v}$ and then solve the matrix equation

$$A\vec{b} = \vec{y} \tag{2}$$

for the unknown vector $\vec{b}$ with the fit coefficients. Since the basis functions

$$g_k(x) = x^{k-1} \tag{3}$$

the matrix A is

$$A = \begin{pmatrix} n & \sum x_i & \cdots & \sum x_i^{p-2} & \sum x_i^{p-1} \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{p-1} & \sum x_i^{p} \\ \vdots & \vdots & \vdots & & \\ \sum x_i^{n-2} & \sum x_i^{n-1} & \cdots & \sum x_i^{n+p-4} & \sum x_i^{n+p-3} \\ \sum x_i^{n-1} & \sum x_i^{n} & \cdots & \sum x^{n+p-3} & \sum x^{n+p-2} \end{pmatrix} \tag{4}$$

$$\vec{y} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^{p-2} y_i \\ \sum x_i^{p-1} y_i \end{pmatrix} \tag{5}$$

The Jacobian needed to find the covariance matrix is

$$J_{i,j} = x_i^{j-1} \tag{6}$$

$$J = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{p-2} & x_1^{p-1} \\ 1 & x_2 & \cdots & x_2^{p-2} & x_2^{p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n-1} & \cdots & x_{n-1}^{p-2} & x_{n-1}^{p-1} \\ 1 & x_n & \cdots & x_n^{p-2} & x_n^{p-1} \end{pmatrix} \tag{7}$$

$$J^T J = \begin{pmatrix} \sum x_i^{p-1} & \sum x_i^{p} & \cdots & \sum x_i^{2p-3} & \sum x_i^{2p-2} \\ \sum x_i^{p-2} & \sum x_i^{p-1} & \cdots & \sum x_i^{2p-4} & \sum x_i^{2p-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{p-1} & \sum x_i^{p} \\ n & \sum x_i & \cdots & \sum x_i^{p-2} & \sum x_i^{p-1} \end{pmatrix} \tag{8}$$

As usual, the uncertainties in the fit parameters will be found by

$$C = (J^T J)^{-21} \tag{9}$$

$$\sigma_i = \sigma_y C_{ii} \tag{10}$$

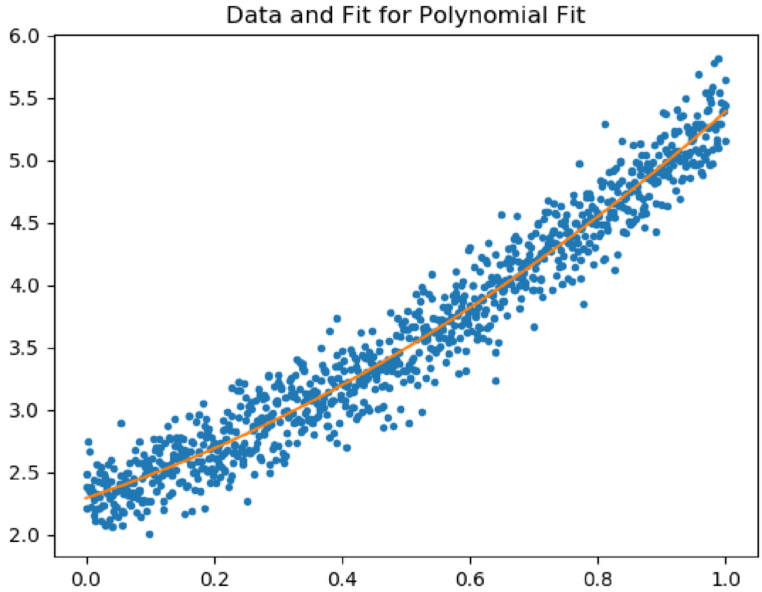

Figure 1: Data (blue dots) and the quadratic polynomial fit to the data (orange line).

3 Using Orthogonal Polynomials

Polynomials as a sum of monomials are not always be best way to fit data. The problem is that the quadratic and linear terms can be highly correlated. Consider fitting the quadratic function

$$y = b_0 + b_1 x + b_2 x^2 \tag{11}$$

for $b_0 = 2.3$, $b_1 = 1.7$ and $b_2 = 1.4$ for $0 \leq x \leq 1$ with random noise having an uncertainty of $sigma_y = 0.2$. I used 1,000 data points.

Fig. 1 is a plot of the data and fit for the polynomial case. The fit does a good job following the data and the fit parameters are good: $b_1 = 2.29$, $b_2 = 1.70$, and $b_3 = 1.40$. However, the covariance matrix presents problems.

$$C = \begin{pmatrix} 3.69e-4 & -1.47e-3 & 1.23e-3 \\ -1.47e-3 & 7.87e-3 & -7.38e-3 \\ 1.23e-3 & -7.38e-3 & 7.38e-3 \end{pmatrix} \tag{12}$$

The off diagonal elements are about the same size as the diagonal elements. This means the fit parameters are highly correlated with each other. This can be illustrated further by finding the eigenvectors of this matrix. Those tell us what linear combinations of monomials are statistically independent. The eigenvectors are the columns of the matrix

$$C' = \begin{pmatrix} -0.128 & 0.827 & -0.548 \\ 0.714 & 0.460 & 0.528 \\ -0.689 & 0.324 & 0.649 \end{pmatrix} \tag{13}$$

Notice that the eigenvectors strongly mix each of the monomial components.

An alternative is to fit the curve to an expansion in terms of the "shifted" Legendre polynomials which are orthogonal on the interval $[0, 1]$.

$$\bar{P}_0(x) = 1 \tag{14}$$

$$\bar{P}_1(x) = 2x - 1 \tag{15}$$

$$\bar{P}_2(x) = 6x^2 - 6x + 1 \tag{16}$$

Fitting these values to the data produces a much nicer covariance matrix.

$$C = \begin{pmatrix} 4.12e-5 & 3.50e-21 & -2.05e-7 \\ 3.5e-21 & 1.23e-4 & -1.75e-23 \\ -2.05e-7 & -1.75e-23 & 2.05e-4 \end{pmatrix} \tag{17}$$

The off-diagonal elements are now much smaller than the diagonal ones. As one would expect in this case, the eigenvectors

$$C' = \begin{pmatrix} -1.000 & -0.000 & -0.001 \\ 0.000 & -1.000 & -0.000 \\ -0.001 & 0.000 & 1.000 \end{pmatrix} \tag{18}$$

also show that these coefficients of these polynomials are statistically independent.

4 Scaling the Data

Whether or not you are using orthogonal polynomials, it is a good idea to express the polynomials in terms of a quantity $0 \le t \le 1$. If $x_i \le x \le x_f$,

$$t = \frac{x - x_i}{x_f - x_i} \tag{19}$$

The coefficients of polynomials in t will be more stable numerically than polynomials in x if $|x| \gg 1$.

5 Regression

A question which arises in polynomial fitting is which terms make sense to keep statistically. A good way to decide is based on the Student t distribution. This gives the probability density that a fit parameter with mean μ and standard error σ will be t

standard deviations away from the origin. Let $T(t, \nu)$ be the Student t distribution. Then

$$p = 2 * \left(1 - \int_0^{\mu} /\sigma T(\tau, \nu)\, d\tau\right) \tag{20}$$

which is called the p value is the probability that a fit parameter value of μ or greater could have happened by chance. The parameter ν is the number of degrees of freedom in the fit. If n is the number of data points and f is the number of fit parameters,

$$\nu = n - f \tag{21}$$

A good standard for a fit parameter to be statistically significant is to have $p < 0.05$.

References

[1] R. Steven Turley, "Linear Least Squares," BYU, 2018.