Polynomial Regression

http://polynomialregression.drque.net/math.html

Minimizing the Residual Error

$$r(x)=\sum_{i=0}^{n}\left[y_i-f(x_i)\right]^2$$

- So now that we have a function that measures residual, what do we do with it? Well, if we are trying to produce a function that models a set of data, we want the residual to be as small as possible—we want to minimize it. This would produce a function that deviates from the observed data as little as possible.
- To minimize, we need to be able to modify the function's coefficients as the coefficients are the variables we have to manipulate for finding the best fit. In our line function $y = m\,x + b$, the coefficients are m and b. There is an added benefit to squaring the residuals—the square of residual forms a parabola.

Minimize the residual error and determining the coefficients..

$$r(x)=\sum_{i=0}^{n}[y_i-f(x_i)]^2$$

- To see why this is useful, consider a 1st degree polynomial with three known points (10, 8, 11). Let's make the function:

$$f(x)=c_0$$

You will notice x isn't used, but this is legitimate. Use this to construct the residual function:

$$r(x)=\sum_{i=0}^{n}[y_i-c_0]^2$$

And expand this for our data:

$$r(x)=\sum_{i=0}^{n}[y_i-c_0]^2=(10-c_0)^2+(8-c_0)^2+(11-c_0)^2$$

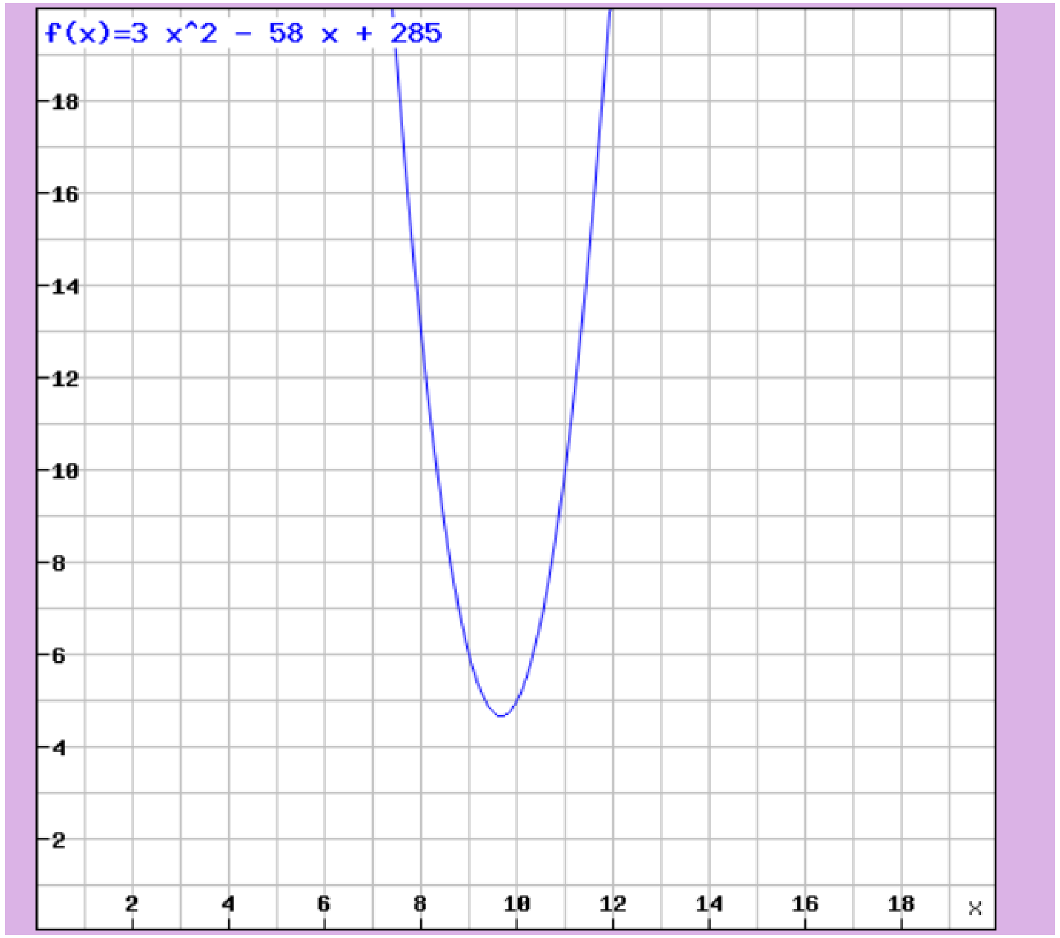
If we multiply everything and collect like terms we end up with:

$$r(x)=3c_0^2-58c_0+285$$

$$r(x) = 3c_0^2 - 58c_0 + 285$$

Now graph this function, but use c_0 as the variable:

- The graph is a parabola. A parabola always has one and only one lowest point—it's vertex. The lowest point is the point with the lowest amount of error. So if we find coefficients that yield the vertex we have the coefficients that produce the lowest error.
- Finding the vertex of a parabola for a function that has a single coefficient is easy. Take the derivative of the function and find where that function is equal to zero. This works because any time the derivative of a function is equal to zero, that point is either a local maximum or minimum. A parabola only has one such point: the minimum.



$$r(x)=3c_0^2-58c_0+285$$

Using our example start by taking the derivative:

$$\frac{dr}{dc_0}=6c_0-58$$

(Solution.)

And set this equal to zero:

$$6c_0-58=0$$

Now solve for c_0:

$$c_0=\frac{58}{6}=9\frac{2}{3}$$

So the coefficient that best fits this data is $c_0 = 9\ 2/3$. Our function is then:

$$f(x)=9\frac{2}{3}$$

We have just solved a 0 degree polynomial using least squares, and our coefficient is the average of our data set:

$$\frac{10+8+11}{3}=9\frac{2}{3}$$

A polynomial of degree 0 is just a constant because $f(x) = c_0\ x^0 = c_0$. Likewise preforming polynomial regression with a degree of 0 on a set of data returns a single constant value. It is the same as the mean average of that data. This makes sense because the average is an approximation of all the data points.

Paraboloid's… for higher orders

- See the example for y = mx + b linear regression
- See the example for ax^2 + bx + c linear regression
- See the generic solution…
- Residual errors…