Least Squares Polynomial Algorithm

December 27, 2013

Joseph Pousada

Often when conducting research one will use the least squares polynomial method to derive a model that closest fits the data. Various statistical software is typically used for this. For those with an inquiring mind that want to know what the software is doing, here is an example on how to compute a third degree least square polynomial manually. (All manual steps were done in Maple.)

So let us assume that you are studying teeth of local alligators. Let's assume that you tag an alligator when it was born and decide to study teeth length/size at 13 months to 18 months. You gather the following data below
$(x\ values = months, y\ values = teeth\ size\ cm)$.

x_i	13	14	15	16	17	18
y_i	1.0	1.003	1.028	1.101	1.201	1.402

So now that we have our data we will show the manual steps in Maple below on deriving a third degree least squares polynomial.

$restart:$
$x1 := 13:$
$x2 := 14:$
$x3 := 15:$
$x4 := 16:$
$x5 := 17:$
$x6 := 18:$

$y1 := 1.0:$
$y2 := 1.003:$
$y3 := 1.028:$
$y4 := 1.101:$
$y5 := 1.201:$
$y6 := 1.402:$

$s0 := ((x1)^0 + (x2)^0 + (x3)^0 + (x4)^0 + (x5)^0 + (x6)^0);$

$$6 \quad (1)$$

$s1 := ((x1)^1 + (x2)^1 + (x3)^1 + (x4)^1 + (x5)^1 + (x6)^1);$

$$93$$

$s2 := ((x1)^2 + (x2)^2 + (x3)^2 + (x4)^2 + (x5)^2 + (x6)^2);$

$$1459 \quad (3)$$

$s3 := ((x1)^3 + (x2)^3 + (x3)^3 + (x4)^3 + (x5)^3 + (x6)^3);$

$$23157 \quad (4)$$

$s4 := ((x1)^4 + (x2)^4 + (x3)^4 + (x4)^4 + (x5)^4 + (x6)^4);$

$$371635 \quad (5)$$

$s5 := ((x1)^5 + (x2)^5 + (x3)^5 + (x4)^5 + (x5)^5 + (x6)^5);$

$$6026493 \quad (6)$$

$s6 := ((x1)^6 + (x2)^6 + (x3)^6 + (x4)^6 + (x5)^6 + (x6)^6);$

$$98673979 \quad (7)$$

$syx0 := ((y1) \cdot (x1)^0 + (y2) \cdot (x2)^0 + (y3) \cdot (x3)^0 + (y4) \cdot (x4)^0 + (y5) \cdot (x5)^0 + (y6) \cdot (x6)^0);$

$$6.735 \quad (8)$$

$syx1 := ((y1) \cdot (x1)^1 + (y2) \cdot (x2)^1 + (y3) \cdot (x3)^1 + (y4) \cdot (x4)^1 + (y5) \cdot (x5)^1 + (y6) \cdot (x6)^1);$

$$105.731 \quad (9)$$

$syx2 := ((y1) \cdot (x1)^2 + (y2) \cdot (x2)^2 + (y3) \cdot (x3)^2 + (y4) \cdot (x4)^2 + (y5) \cdot (x5)^2 + (y6) \cdot (x6)^2);$

$$1680.081 \quad (10)$$

$syx3 := ((y1) \cdot (x1)^3 + (y2) \cdot (x2)^3 + (y3) \cdot (x3)^3 + (y4) \cdot (x4)^3 + (y5) \cdot (x5)^3 + (y6) \cdot (x6)^3);$

$$27005.405 \quad (11)$$

Getting our equations:

$eq1 := a0 \cdot s0 + a1 \cdot s1 + a2 \cdot s2 + a3 \cdot s3 = syx0;$

$$6\, a0 + 93\, a1 + 1459\, a2 + 23157\, a3 = 6.735 \quad (12)$$

$eq2 := a0 \cdot s1 + a1 \cdot s2 + a2 \cdot s3 + a3 \cdot s4 = syx1;$

$$93\, a0 + 1459\, a1 + 23157\, a2 + 371635\, a3 = 105.731 \quad (13)$$

$eq3 := a0 \cdot s2 + a1 \cdot s3 + a2 \cdot s4 + a3 \cdot s5 = syx2;$

$$1459\, a0 + 23157\, a1 + 371635\, a2 + 6026493\, a3 = 1680.081 \quad (14)$$

$eq4 := a0 \cdot s3 + a1 \cdot s4 + a2 \cdot s5 + a3 \cdot s6 = syx3;$

$$23157\, a0 + 371635\, a1 + 6026493\, a2 + 98673979\, a3 = 27005.405 \quad (15)$$

$solve(\{eq1, eq2, eq3, eq4\}, \{a0, a1, a2, a3\});$

$$\{a0 = -5.802730159, a1 = 1.562493386, a2 = -0.1199087302, a3 = 0.003074074074\} \quad (16)$$

So our polynomial is:

$P_3(x) = (-5.802730159) + (1.562493386) \cdot x + (-0.1199087302) \cdot x^2 + (0.003074074074) \cdot x^3$

Computing the projected values with our polynomials and the errors:

$P3 := x \rightarrow (-5.802730159) + (1.562493386) \cdot x + (-0.1199087302) \cdot x^2 + (0.003074074074) \cdot x^3;$

$$x \rightarrow -5.802730159 + 1.562493386\, x + (-1) \cdot 0.1199087302\, x^2 + 0.003074074074\, x^3 \quad (17)$$

$p1 := P3(x1);$

$$0.998849201 \quad (18)$$

$p2 := P3(x2);$

$$1.005325379 \quad (19)$$

$p3 := P3(x3);$

$$1.03020633 \quad (20)$$

$p4 := P3(x4);$

$$1.09193650 \quad (21)$$

$p5 := P3(x5);$

$$1.20896030 \quad (22)$$

$p6 := P3(x6);$

$$1.39972221 \quad (23)$$

Computing the error...

$e1 := y1 - P3(x1);$

$$0.001150799 \quad (24)$$

$e2 := y2 - P3(x2);$

$$-0.002325379 \quad (25)$$

$e3 := y3 - P3(x3);$

$$-0.00220633 \quad (26)$$

$e4 := y4 - P3(x4);$

$$0.00906350 \quad (27)$$

$e5 := y5 - P3(x5);$

$$-0.00796030 \quad (28)$$

$e6 := y6 - P3(x6);$

$$-0.001556019 \quad (29)$$

So our chart becomes:

i	1	2	3	4	5	6
x_i	13	14	15	16	17	18
y_i	1.0	1.003	1.028	1.101	1.201	1.402
$P(x_i)$	0.998849201	1.005325379	1.03020633	1.09193650	1.20896030	1.39972221
$y_i - P(x_i)$	0.001150799	−0.002325379	-0.00220633	0.00906350	−0.00796030	−0.001556019