Notes on fitting polynomials

Given pairs of data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ consider approximating polynomials of the form $p(t) = a_0 + a_1 x + a_2 x^2 + \cdots + a_t x^t$. The error e_i for the i^{th} pair is the distance between y_i and the height $y(x_i)$ of the polynomial at x_i. This is $e_i = y_i - (a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_t x_i^t)$. If we consider the equations $a_0 + a_1 x + a_2 x^2 + \cdots + a_t x^t = y_i$ for $i = 1, 2, \ldots, n$ in the variables $a_0, a_1, \ldots, a_t$ we can think of this as a system of equations $A\boldsymbol{x} = \boldsymbol{b}$ where $A = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^t \\ 1 & x_2 & x_2^2 & \ldots & x_2^t \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^t \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_t \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$. We can then use the best approximating solutions in any of the norms we have discussed before.

In the special case that the x_i are distinct and $t + 1 = n$ we will not need to approximate. There will be an exact solution. This extend the idea that two points determine a line to n points (with distinct x coordinates) determine a unique polynomial of degree $n - 1$. This follows because the matrix in this case (called the VanderMonde matrix) can be shown to have an inverse so that we can solve $A\boldsymbol{x} = \boldsymbol{b}$ uniquely as $A^{-1}\boldsymbol{b}$.

Rather than working out A^{-1} explicitly we get the polynomial using Lagrange interpolating polynomials. Let $L_k(x) = \frac{(x-x_1)\ldots(x-x_{k-1})(x-x_{k+1})\ldots(x-x_n)}{(x_k-x_1)\ldots(x_k-x_{k-1})(x_k-x_{k+1})\ldots(x_k-x_n)}$. Note that the x_i are given values and x is a variable. This is a polynomial with degree $n - 1$. The factors in the numerator show that the roots are $x_1, x_2, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$. Note also that substituting x_k for x in the numerator we get the denominator. So we have $L_k(x_i) = 0$ for $i \not k$ and $L_k(x_k) = 1$. Then $p(x) = y_1 L_1(x) + y_2 L_2(x) + \cdots + y_n L_n(x)$ satisfies $p(x_i) = y_i$ for $i = 1, 2, \ldots, n$. So this polynomial goes through all of the points (and from the comments above it is unique).

For example, if we want to find the unique parabola (degree 2 polynomial) through the points $(1, 4), (2, 3), (3, 8)$ we could solve the equations $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 8 \end{bmatrix}$ to determine the coefficients a_0, a_1, a_2 as $a_0 = 11, a_1 = -10, a_2 = 3$.

Alternatively we have $L_1(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{1}{2}(x^2 - 5x + 6)$ and $L_2 = \frac{(x-1)(x-3)}{(2-1)(2-3)} = (-1)(x^2 - 4x + 3)$ and $L_3(x) = \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{1}{2}(x^2 - 3x + 2)$. Then $p(x) = 4\,\frac{1}{2}(x^2 - 5x + 6) + 3\,(-1)(x^2 - 4x + 3) + 8\,\frac{1}{2}(x^2 - 3x + 2) = 3x^2 - 10x + 11$.