This note will explain the method of least squares applied to polynomial approximation. Suppose we have \( m + 1 \) data points \((x_j, y_j), \ j = 0, \ldots, m\) and want to find a polynomial fit to the data. One way is to use method of least squares. Let the polynomial be expressed as 
\[
a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.
\]
We consider the equations
\[
a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n = y_0
\]
\[
a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n = y_1
\]
\[
\vdots
\]
\[
a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_n x_m^n = y_m.
\]
Rewrite these equations as
\[
\begin{bmatrix}
1 & x_0 & x_0^2 & \cdots & x_0^n \\
1 & x_1 & x_1^2 & \cdots & x_1^n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_m & x_m^2 & \cdots & x_m^n
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_n
\end{bmatrix}
= 
\begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_m
\end{bmatrix}.
\]
Let \( X \) denote
\[
X = \begin{bmatrix}
1 & x_0 & x_0^2 & \cdots & x_0^n \\
1 & x_1 & x_1^2 & \cdots & x_1^n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_m & x_m^2 & \cdots & x_m^n
\end{bmatrix},
\]
and
\[
a = \begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_n
\end{bmatrix}, \quad y = \begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_m
\end{bmatrix}.
\]
Then the normal equations are
\[
X^T X a = X^T y.
\]
This can be written symbolically as
\[
\begin{bmatrix}
\sum x & \sum x^2 & \cdots & \sum x^n \\
\sum x & \sum x^2 & \cdots & \sum x^{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
\sum x^n & \sum x^{n+1} & \cdots & \sum x^{2n}
\end{bmatrix}
\begin{bmatrix}
a \\
a \\
\vdots \\
a
\end{bmatrix}
= 
\begin{bmatrix}
\sum y \\
\sum xy \\
\vdots \\
\sum x^n y
\end{bmatrix}.
If all $x_i$ are distinct and $m \geq n$ then $X$ is nonsingular. The proof is as follows. Suppose there is a vector $a$ so that $Xa = 0$. Then the polynomial $a_0 + a_1x + a_2x^2 + \ldots a_nx^n$ vanishes at $m + 1$ distinct points $x_0, x_1, x_2, \ldots, x_m$. Since $m \geq n$, the polynomial is the zero polynomial, i.e. $a = 0$.

This implies that $X^TX$ is a square nonsingular matrix. For if $X^TXa = 0$ then $a^TX^TXa = \|Xa\|^2 = 0$. Hence $Xa = 0$, so $a = 0$. 

\textit{least squares}