Discrete Least Squares Approximation

MATH 375 *Numerical Analysis*

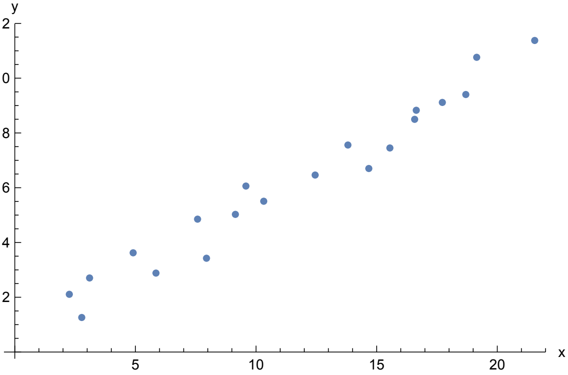
J Robert Buchanan

Department of Mathematics

Spring 2022

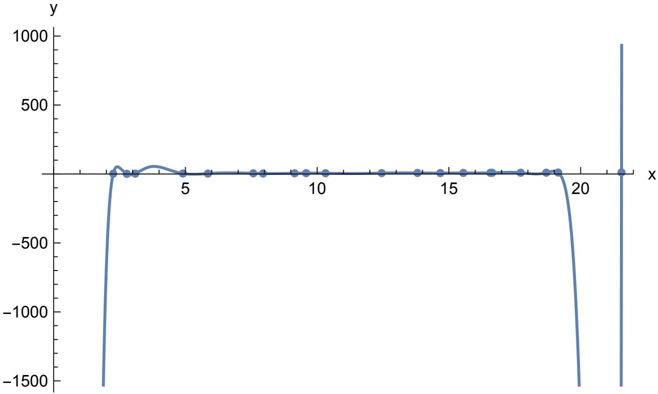
Background

A task which researchers frequently face is that of finding the parameters of a “simple” function so that its graph passes in some sense “close” to a scatter plot of data organized as ordered pairs.



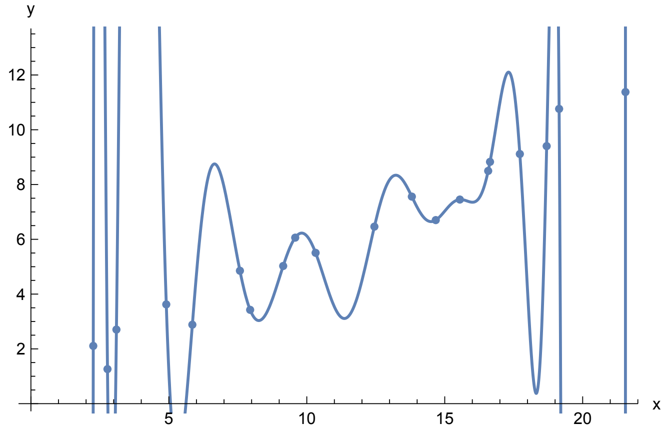
Lagrange Interpolating Polynomial

Consider the graph of the Lagrange interpolating polynomial fit to the data.



Lagrange Interpolating Polynomial

Consider the graph of the Lagrange interpolating polynomial fit to the data.



Notion of “Close”

Since the data points were roughly arrayed linearly in the *xy*-plane, perhaps we should find the equation of a line which is “close” to the data points.

Notion of “Close”

Since the data points were roughly arrayed linearly in the *xy*-plane, perhaps we should find the equation of a line which is “close” to the data points.

Q: what sense of “close”-ness do we want to use?

Notion of "Close"

Since the data points were roughly arrayed linearly in the *xy*-plane, perhaps we should find the equation of a line which is "close" to the data points.

Q: what sense of "close"-ness do we want to use?

- **Minimax problem**: minimize

$$\max_{i=1,\ldots,n} \left\{ |y_i - (ax_i + b)| \right\}.$$

- **Absolute deviation problem**: minimize

$$\sum_{i=1}^{n} |y_i - (ax_i + b)|.$$

- **Least squares problem**: minimize

$$\sum_{i=1}^{n} (y_i - (ax_i + b))^2.$$

Linear Least Squares (1 of 3)

Define $f(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$, then a necessary condition for the attainment of a minimum is that a and b satisfy the simultaneous equations:

$$\begin{aligned} \frac{\partial f}{\partial a} &= 0 \\ \frac{\partial f}{\partial b} &= 0. \end{aligned}$$

Linear Least Squares (1 of 3)

Define $f(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$, then a necessary condition for the attainment of a minimum is that a and b satisfy the simultaneous equations:

$$\begin{aligned} \frac{\partial f}{\partial a} &= 0 \\ \frac{\partial f}{\partial b} &= 0. \end{aligned}$$

Upon differentiating, these equations take the form:

$$\begin{aligned} 0 &= -2 \sum_{i=1}^{n} x_i (y_i - (ax_i + b)) \\ 0 &= -2 \sum_{i=1}^{n} (y_i - (ax_i + b)). \end{aligned}$$

Linear Least Squares (2 of 3)

$$
\begin{aligned}
0 &= -2\sum_{i=1}^{n} x_i(y_i - (ax_i + b)) \\
0 &= -2\sum_{i=1}^{n} (y_i - (ax_i + b)).
\end{aligned}
$$

Linear Least Squares (2 of 3)

$$
\begin{aligned}
0 &= -2\sum_{i=1}^{n} x_i(y_i - (ax_i + b)) \\
0 &= -2\sum_{i=1}^{n} (y_i - (ax_i + b)).
\end{aligned}
$$

The terms in these equations can be re-arranged to yield the **normal equations**:

$$
\begin{aligned}
a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} x_i y_i \\
a\sum_{i=1}^{n} x_i + b\,n &= \sum_{i=1}^{n} y_i.
\end{aligned}
$$

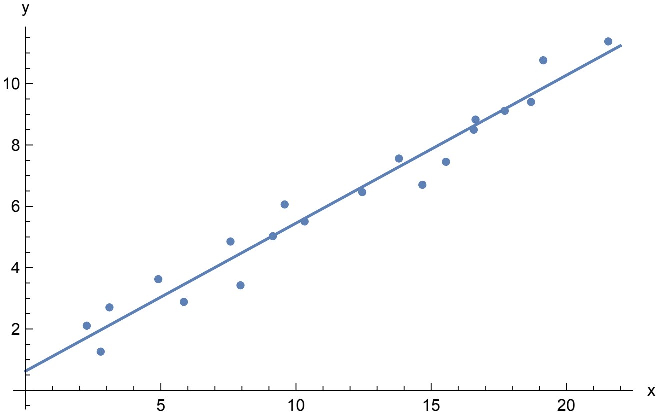
Linear Least Squares (3 of 3)

Solving for a and b produces:

$$
\begin{aligned}
a &= \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \\
b &= \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.
\end{aligned}
$$

Linear Fit

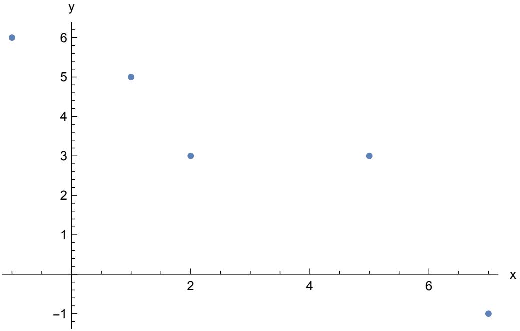
For the data in our sample, $a = 0.482151$ and $b = 0.629109$.



Example (1 of 4)

Find the linear least squares best fitting line for the data in the following table.

x	y
-1	6
1	5
2	3
5	3
7	-1



Example (2 of 4)

The following table can help us organize our hand calculations.

x_i	y_i	x_i^2	$x_i y_i$
-1	6	1	-6
1	5	1	5
2	3	4	6
5	3	25	15
7	-1	49	-7
$\sum_{i=1}^{5} x_i = 14$	$\sum_{i=1}^{5} y_i = 16$	$\sum_{i=1}^{5} x_i^2 = 80$	$\sum_{i=1}^{5} x_i y_i = 13$

Example (3 of 4)

$$\sum_{i=1}^{5} x_i = 14$$

$$\sum_{i=1}^{5} y_i = 16$$

$$\sum_{i=1}^{5} x_i^2 = 80$$

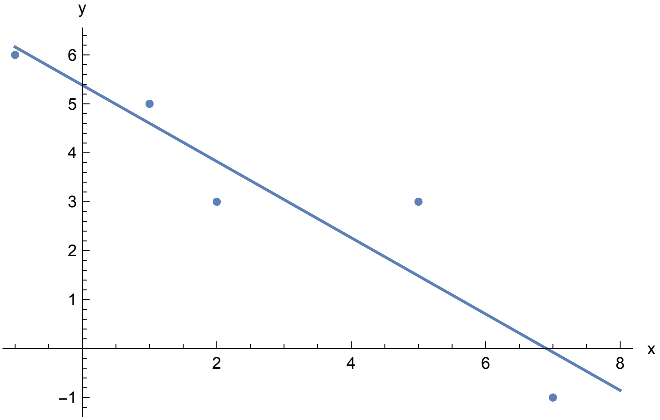
$$\sum_{i=1}^{5} x_i y_i = 13$$

$$a = \frac{5(13) - (14)(16)}{5(80) - (14)^2} = -\frac{53}{68}$$

$$b = \frac{(80)(16) - (13)(14)}{5(80) - (14)^2} = \frac{183}{34}$$

Example (4 of 4)

$$y = -\frac{53}{68}x + \frac{183}{34}$$



Exponential Model

(x, y) may be suspected of having an exponential relationship of one of the two forms:

$$
\begin{aligned}
y &= b\,e^{a\,x} \quad \text{or} \\
y &= b\,x^{a}.
\end{aligned}
$$

If we take the logarithm of both sides of these equations we regain a linear equation.

Exponential Model

(x, y) may be suspected of having an exponential relationship of one of the two forms:

$$
\begin{aligned}
y &= b\,e^{a\,x} \quad \text{or} \\
y &= b\,x^{a}.
\end{aligned}
$$

If we take the logarithm of both sides of these equations we regain a linear equation.

$$
\begin{aligned}
\ln y &= \ln b + a\,x \quad \text{or} \\
\ln y &= \ln b + a \ln x
\end{aligned}
$$

In both cases we can use linear least squares to determine a and b.

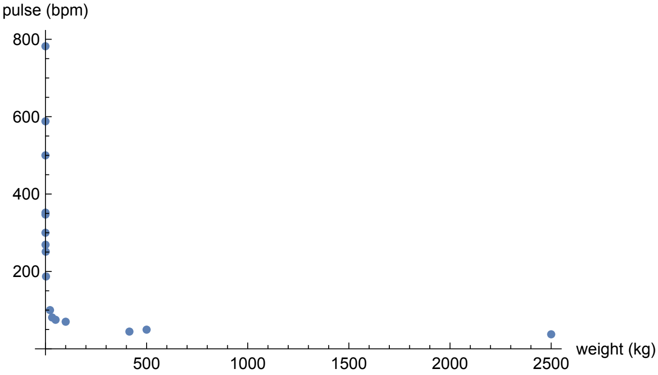
Pulse Rate *vs.* Body Mass (1 of 6)

Consider the following data collected from a range of mammals.

Mass (kg)	Pulse (bpm)	Mass (kg)	Pulse (bpm)
0.0035	782	2.7	187
0.006	588	22.5	100
0.017	500	33	81
0.103	347	50	75
0.117	300	100	70
0.252	352	415	44.5
0.437	269	500	49.5
1.34	251	2500	37.5

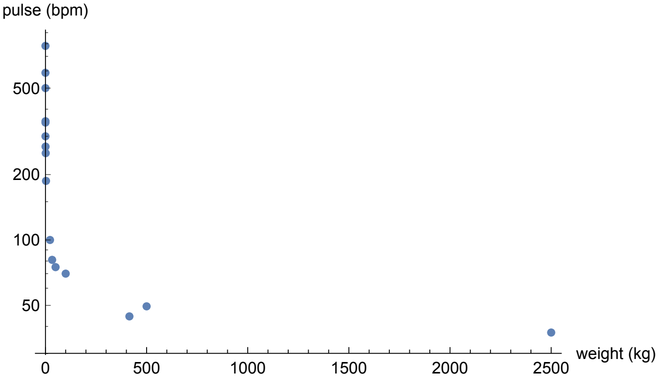
Pulse Rate *vs.* Body Mass (2 of 6)

Scatter plot of the data.



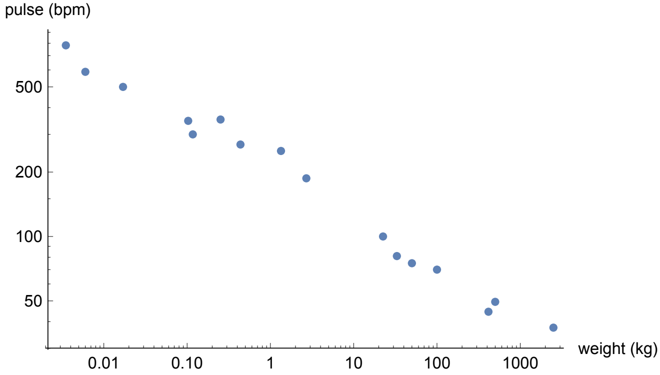
Pulse Rate *vs.* Body Mass (3 of 6)

Suppose we assume $p = b\,e^{a\,w}$, then a log plot resembles the following.



Pulse Rate *vs.* Body Mass (4 of 6)

Suppose we assume $p = b\,w^a$, then a log-log plot resembles the following.



Pulse Rate *vs.* Body Mass (5 of 6)

$$\begin{aligned} p &= b\,w^a \\ \ln p &= \ln b + a \ln w \end{aligned}$$

Using the linear least squares formulas we find:

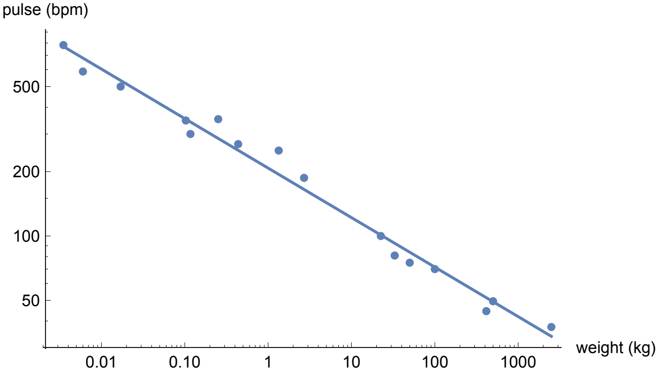
$$\begin{aligned} a &= -0.231649 \\ \ln b &= 5.3366 \\ b &= 207.805 \end{aligned}$$

and thus

$$p = 207.805 w^{-0.231649}.$$

Pulse Rate *vs.* Body Mass (6 of 6)

$$p = 207.805 w^{-0.231649}.$$



General Case

Suppose we wish to fit a polynomial of degree n to some data.

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

There are $n + 1$ coefficients to be estimated.

General Case

Suppose we wish to fit a polynomial of degree n to some data.

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

There are $n + 1$ coefficients to be estimated.

Define the function

$$f(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{m} (y_i - p(x_i))^2 .$$

A necessary condition for f to be minimized is

$$\begin{aligned} \frac{\partial}{\partial a_0}[f(a_0, a_1, \ldots, a_n)] &= 0 \\ \frac{\partial}{\partial a_1}[f(a_0, a_1, \ldots, a_n)] &= 0 \\ &\vdots \\ \frac{\partial}{\partial a_n}[f(a_0, a_1, \ldots, a_n)] &= 0. \end{aligned}$$

Calculating Partial Derivatives

$$
\begin{aligned}
0 &= -2\sum_{i=1}^{m}(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n) \\
0 &= -2\sum_{i=1}^{m} x_i(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n) \\
&\vdots \\
0 &= -2\sum_{i=1}^{m} x_i^j(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n) \\
&\vdots \\
0 &= -2\sum_{i=1}^{m} x_i^n(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n)
\end{aligned}
$$

Normal Equations

$$
\begin{aligned}
a_0 m + a_1 \sum_{i=1}^{m} x_i + \cdots + a_n \sum_{i=1}^{m} x_i^n &= \sum_{i=1}^{m} y_i \\
a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 + \cdots + a_n \sum_{i=1}^{m} x_i^{n+1} &= \sum_{i=1}^{m} x_i y_i \\
&\vdots \\
a_0 \sum_{i=1}^{m} x_i^j + a_1 \sum_{i=1}^{m} x_i^{1+j} + \cdots + a_n \sum_{i=1}^{m} x_i^{n+j} &= \sum_{i=1}^{m} x_i^j y_i \\
&\vdots \\
a_0 \sum_{i=1}^{m} x_i^n + a_1 \sum_{i=1}^{m} x_i^{n+1} + \cdots + a_n \sum_{i=1}^{m} x_i^{2n} &= \sum_{i=1}^{m} x_i^n y_i
\end{aligned}
$$

Solving the Normal Equations

If all the x_i's are distinct then the system of normal equations has a unique solution.

Solving the Normal Equations

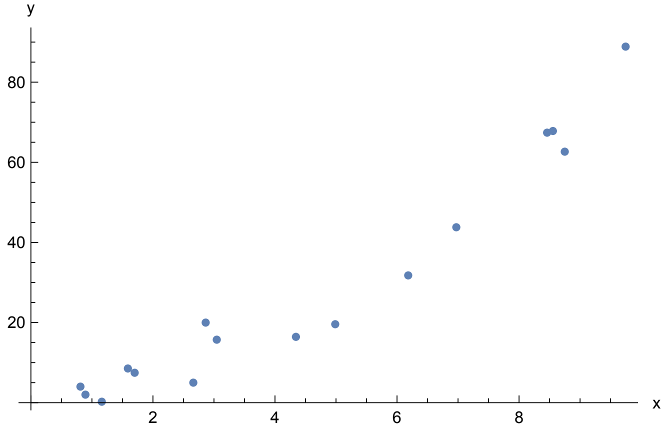
If all the x_i's are distinct then the system of normal equations has a unique solution.

Example

Find the best fitting quadratic polynomial for the data in the table below.

x	y	x	y
0.811716	4.00967	4.34336	16.4355
0.893654	2.0062	4.98746	19.5864
1.1615	0.242184	6.1835	31.7715
1.58907	8.53801	6.97181	43.7873
1.70058	7.47043	8.45983	67.3983
2.66158	5.0021	8.55456	67.8103
2.86422	19.9925	8.74913	62.6538
3.046	15.7271	9.74663	88.8716

Plot of the Data



Example: Normal Equations

If $p(x) = a_0 + a_1 x + a_2 x^2$ then

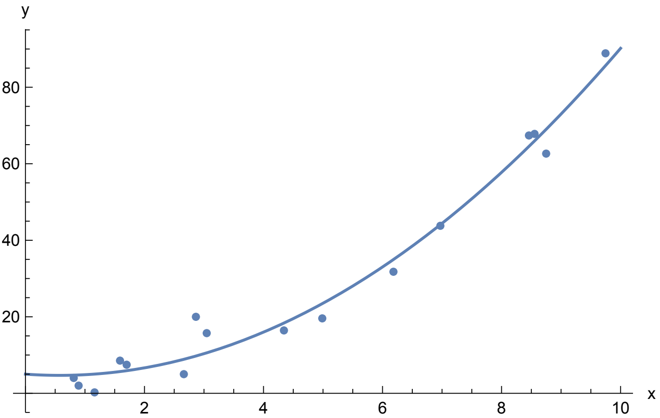
$$\begin{aligned}
16a_0 + 72.7246a_1 + 479.664a_2 &= 461.303 \\
72.7246a_0 + 479.664a_1 + 3690.77a_2 &= 3385.52 \\
479.664a_0 + 3690.77a_1 + 30381.7a_2 &= 27558.0
\end{aligned}$$

which has solution

$$\begin{aligned}
a_0 &= 5.0064 \\
a_1 &= -1.10482 \\
a_2 &= 0.962232 \\
p(x) &= 5.0064 - 1.10482x + 0.962232x^2.
\end{aligned}$$

Fitted Polynomial

$$p(x) = 5.0064 - 1.10482x + 0.962232x^2$$



Homework

- Read Section 8.1.
- Exercises: 1, 3, 6, 7, 13ab