16. Least Squares I. May 13, 2013

16.1. Linear Fit

Assume we conduct some experiment. We input the magnitude \( t \) and measure the output \( y \), and we have:

\[
\begin{array}{c|ccc}
  t & 0 & 1 & 2 \\
y & 0 & 0 & 1 \\
\end{array}
\]

We would like to approximately find \( y(t) \) in the form \( \alpha t + \beta \), where \( \alpha, \beta \) are some coefficients. We cannot fit this linear function precisely into the data, because the points \((0, 0), (1, 0), (2, 1)\) do not lie on the same line. In other words, the system

\[
\begin{align*}
  \alpha 0 + \beta &= 0 \\
  \alpha 1 + \beta &= 0 \\
  \alpha 2 + \beta &= 1 \\
\end{align*}
\]

does not have a solution. But there is a measurement error; so the differences between precise values (left-hand sides) and experimental values (right-hand sides) may be considered as errors. We would like to find \( \alpha, \beta \) so that these errors are minimal. Let us take the sum of squares of these errors and make it minimal:

\[
(\alpha 0 + \beta - 0)^2 + (\alpha 1 + \beta - 0)^2 + (\alpha 2 + \beta - 1)^2 \rightarrow \min.
\]

We can solve this analytically, by taking partial derivatives of the function. But there is a geometric way to do this. The sum of the squares is the distance squared between the vectors

\[
\begin{bmatrix}
  \alpha 0 + \beta \\
  \alpha 1 + \beta \\
  \alpha 2 + \beta
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
  0 \\
  0 \\
  1
\end{bmatrix}
\]

of exact and approximate values of \( y \). Note that

\[
\begin{bmatrix}
  \alpha 0 + \beta \\
  \alpha 1 + \beta \\
  \alpha 2 + \beta
\end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \alpha v_1 + \beta v_2,
\]

so this is the plane generated by \( v_1, v_2 \). We should find the values of the parameters \( \alpha, \beta \) such that the vector \( b \) is closest to \( \alpha v_1 + \beta v_2 \). This will be when \( b - \alpha v_1 - \beta v_2 \) is orthogonal to this plane, i.e. when it is orthogonal to both vectors \( v_1, v_2 \). Because we want the vector \( \alpha v_1 + \beta v_2 \) to be the orthogonal projection of \( b \) onto this plane, so the difference between these vectors should be perpendicular to the plane.

So we have:

\[
(b - \alpha v_1 - \beta v_2) \cdot v_1 = 0, \quad (b - \alpha v_1 - \beta v_2) \cdot v_2 = 0.
\]

Rewrite this as

\[
v_1 \cdot \alpha v_1 + v_1 \cdot v_2 \beta = v_1 \cdot b, \quad v_2 \cdot \alpha v_1 + v_2 \cdot v_2 \beta = v_2 \cdot b.
\]

So we have the following system:

\[
\begin{bmatrix}
  v_1 \cdot v_1 & v_1 \cdot v_2 \\
  v_2 \cdot v_1 & v_2 \cdot v_2
\end{bmatrix}
\begin{bmatrix}
  \alpha \\
  \beta
\end{bmatrix} =
\begin{bmatrix}
  b \cdot v_1 \\
  b \cdot v_2
\end{bmatrix}
Plug in the numbers:
\[
\begin{bmatrix}
5 & 3 \\
3 & 3
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}
= \begin{bmatrix}
2 \\
1
\end{bmatrix}
\Rightarrow \alpha = 1/2, \ \beta = -1/6.
\]

So we have:
\[
y(t) = \frac{1}{2} t - \frac{1}{6}
\]

This is the least-squares linear fit to data.

16.2. Quadratic Fit

Assume we have:
\[
\begin{array}{c|ccc}
t & 1 & 0.5 & 2 \\
y & 1 & 2 & 3
\end{array}
\]

We would like to find \( y(t) = at^2 + \beta t + \gamma \). Let us make the system
\[
\begin{align*}
\alpha t^2 + \beta 1 + \gamma &= 1 \\
\alpha 0.5^2 + \beta 0.5 + \gamma &= 2 \\
\alpha t^2 + \beta t + \gamma &= 2 \\
\alpha (-1)^2 + \beta (-1) + \gamma &= 3
\end{align*}
\]

and solve it using least squares method: minimize the distance between
\[
b = \begin{bmatrix}
1 \\
2 \\
2 \\
3
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
\alpha t^2 + \beta 1 + \gamma \\
\alpha 0.5^2 + \beta 0.5 + \gamma \\
\alpha t^2 + \beta t + \gamma \\
\alpha (-1)^2 + \beta (-1) + \gamma
\end{bmatrix}
= \alpha v_1 + \beta v_2 + \gamma v_3,
\]

where
\[
v_1 = \begin{bmatrix}
1 \\
1/4 \\
1
\end{bmatrix}, \quad v_2 = \begin{bmatrix}
1 \\
1/2 \\
2
\end{bmatrix}, \quad v_3 = \begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix}
\]

Do it in the same way:
\[
\begin{bmatrix}
v_1 \cdot v_1 & v_1 \cdot v_2 & v_1 \cdot v_3 \\
v_2 \cdot v_1 & v_2 \cdot v_2 & v_2 \cdot v_3 \\
v_3 \cdot v_1 & v_3 \cdot v_2 & v_3 \cdot v_3
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta \\
\gamma
\end{bmatrix}
= \begin{bmatrix}
b \cdot v_1 \\
b \cdot v_2 \\
b \cdot v_3
\end{bmatrix}
\]

Then plug in the numbers and solve the system.