If \( x \in R^m, m > 1 \), then

\[
f(x) = a^T x + b, \ a \in R^m
\]

Let

\[
E = \sum_{i=1}^{n} (y_i - (a^T x_i + b))^2
\]
Then

\[ \min_{a,b} \sum_{i=1}^{n} (y_i - (a^T x_i + b))^2 \]

\[ \equiv \min_{a,b} \sum_{i=1}^{n} y_i^2 - 2y_i(a^T x_i + b) + (a^T x_i + b)^2 \]

\[ \equiv \min_{a,b} \sum_{i=1}^{n} -2y_i(a^T x_i + b) + (a^T x_i)^2 + 2ba^T x_i + b^2 \]

\[ \equiv \min_{a,b} (\sum_{i=1}^{n} -2y_i x_i)^T a + (\sum_{i=1}^{n} -2y_i) b + \sum_{i=1}^{n} (a^T x_i)^2 + (\sum_{i=1}^{n} 2x_i)^T ab + nb^2 \]
First derivative

\[
\frac{\partial E}{\partial a} = 0 \Rightarrow \left( \sum_{i=1}^{n} -2y_{i}x_{i} \right) + 2 \sum_{i=1}^{n} (a^T x_{i})x_{i} + \left( \sum_{i=1}^{n} 2x_{i} \right) b = 0
\]

\[
\frac{\partial E}{\partial b} = 0 \Rightarrow \left( \sum_{i=1}^{n} -2y_{i} \right) + \left( \sum_{i=1}^{n} 2x_{i} \right)^T a + 2nb = 0
\]
Note that

\[ \sum_{i=1}^{n} (a^T x_i) x_i = \sum_{i=1}^{n} x_i x_i^T a, \]

where

\[ x_i x_i^T \]

is an \( m \) by \( m \) matrix
Data in Higher Dimensional Space V

- Define

\[
S_{xx} = \sum_{i=1}^{n} x_i x_i^T, \quad S_x = \sum_{i=1}^{n} x_i
\]

\[
S_{xy} = \sum_{i=1}^{n} y_i x_i, \quad S_y = \sum_{i=1}^{n} y_i
\]

- Solve a linear system of \(a\) and \(b\)

\[
\begin{bmatrix}
S_{xx} & S_x \\
S_x^T & n
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix}
= \begin{bmatrix}
S_{xy} \\
S_y
\end{bmatrix}
\]

- This system has \(m + 1\) variables
Recall that we solved

$$\min_{a,b} \sum_{i=1}^{n} (y_i - (a^T x_i + b))^2$$

We have

$$a^T x_i + b = \begin{bmatrix} a^T & b \end{bmatrix} \begin{bmatrix} x_i \\ 1 \end{bmatrix}$$

We can consider that each input vector has one extra dimension and the value at that dimension is always 1
We let

$$\bar{a} \leftarrow \begin{bmatrix} a \\ b \end{bmatrix}, \bar{x}_i \leftarrow \begin{bmatrix} x_i \\ 1 \end{bmatrix}$$

The linear system becomes

$$S\bar{x}\bar{a} = S\bar{x}y$$
A Simpler Derivation III

We have

\[ S_{\bar{x}\bar{x}} = \sum_{i=1}^{n} \bar{x}_i \bar{x}_i^T = \sum_{i=1}^{n} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \begin{bmatrix} x_i^T & 1 \end{bmatrix} = \begin{bmatrix} S_{xx} & S_x \\ S_x^T & n \end{bmatrix}, \]

which is the matrix obtained earlier
Because each \( \bar{x}_i \bar{x}_i^T \) is positive semi-definite, we easily get that \( S_{\bar{x}\bar{x}} \) is positive semi-definite.