Multivariate Regression models

Dr. Pritam Anand.
Assistant Professor,
DA-IICT, Gandhinagar.

Multiple Linear Regression model

x_1 Age	x_2 Income (hundred thousand dollar)	y Balance (thousand dollar)
32	0.550798	5.651202
22	0.708148	7.321263
45	0.290905	5.167304
78	0.510828	5.609367
54	0.892947	9.406379
39	0.896293	9.379439
42	0.125585	2.734997
51	0.207243	4.876649
21	0.051467	3.584138
19	0.44081	5.437239

Linear Regression model with two variables

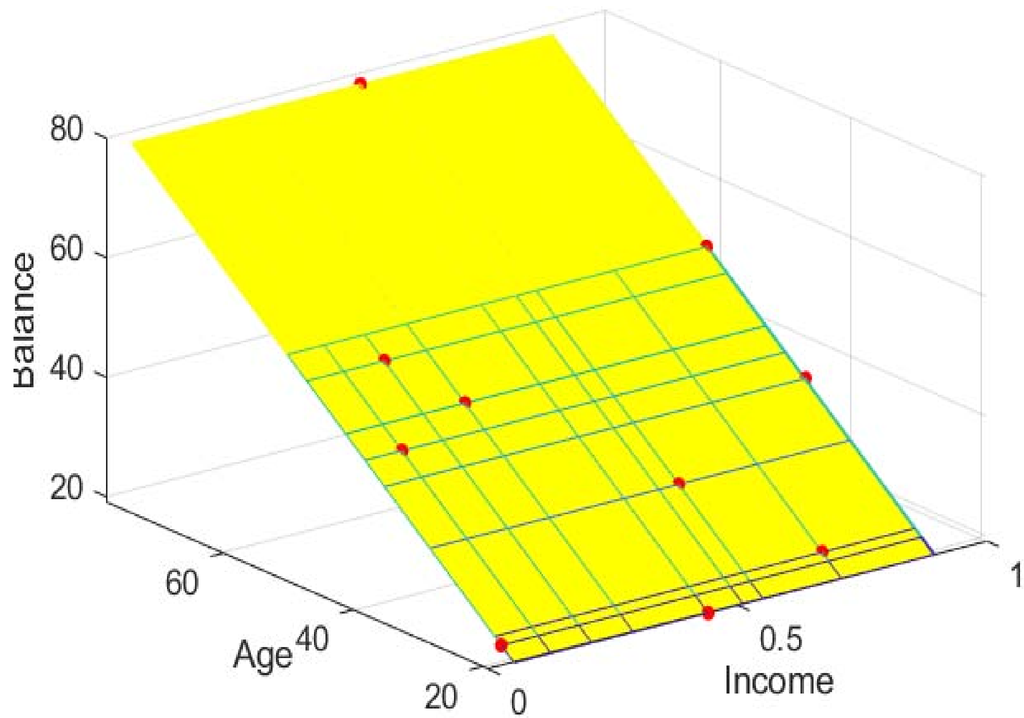
$f(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

Where $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

x_1 Age	x_2 Income (hundred thousand dollar)	y Balance (thousand dollar)
32	0.550798	5.651202
22	0.708148	7.321263
45	0.290905	5.167304
78	0.510828	5.609367
54	0.892947	9.406379
39	0.896293	9.379439
42	0.125585	2.734997
51	0.207243	4.876649
21	0.051467	3.584138
19	0.44081	5.437239

Multiple Linear Regression model

x_1 Age	x_2 Income (hundred thousand dollar)	y Balance (thousand dollar)
32	0.550798	5.651202
22	0.708148	7.321263
45	0.290905	5.167304
78	0.510828	5.609367
54	0.892947	9.406379
39	0.896293	9.379439
42	0.125585	2.734997
51	0.207243	4.876649
21	0.051467	3.584138
19	0.44081	5.437239



For given Training Set T = { (x_{11}, $x_{12,}$,y_1), (x_{21},$x_{22,}$ y_2),....., (x_{n1},$x_{n2,}$ y_n)}, we need to solve

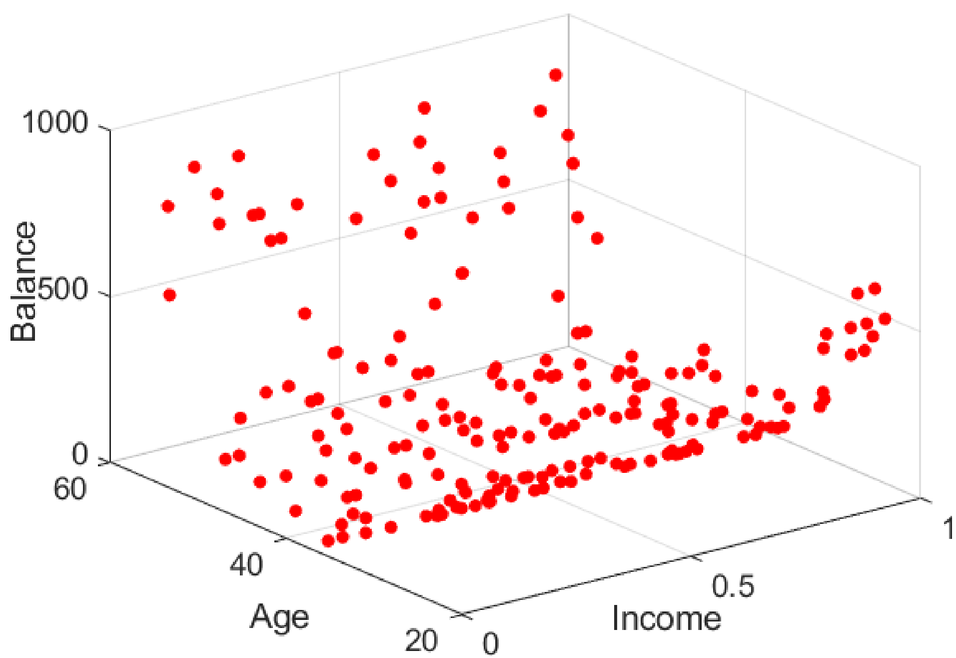
$$Min \quad J(\beta_5,\beta_{4,}\beta_{3,}\beta_{2,}\beta_{1,}\beta_0) = \quad \sum_{i=1}^{n} (y_i-(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))^2 \qquad .....(1)$$

$$u = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} \qquad A = \begin{bmatrix} x_{12} & x_{11} & 1 \\ x_{22} & x_{21} & 1 \\ & & \\ x_{n2} & x_{n1} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ y_n \end{bmatrix}$$

$$\boxed{u = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} = (A^T A)^{-1} \ A^T Y}$$

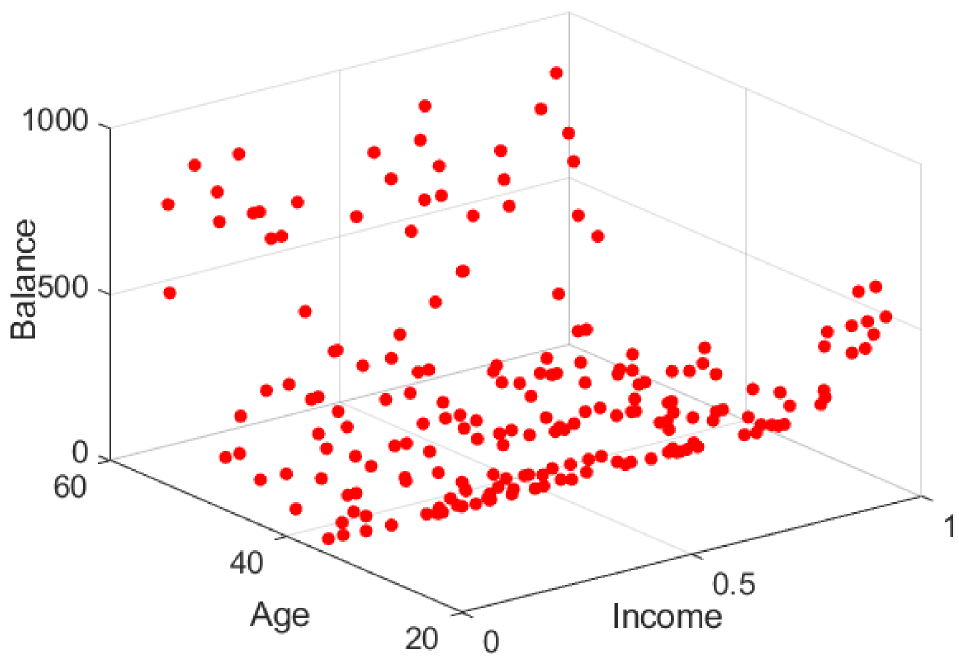
The Least Square problem reduces to

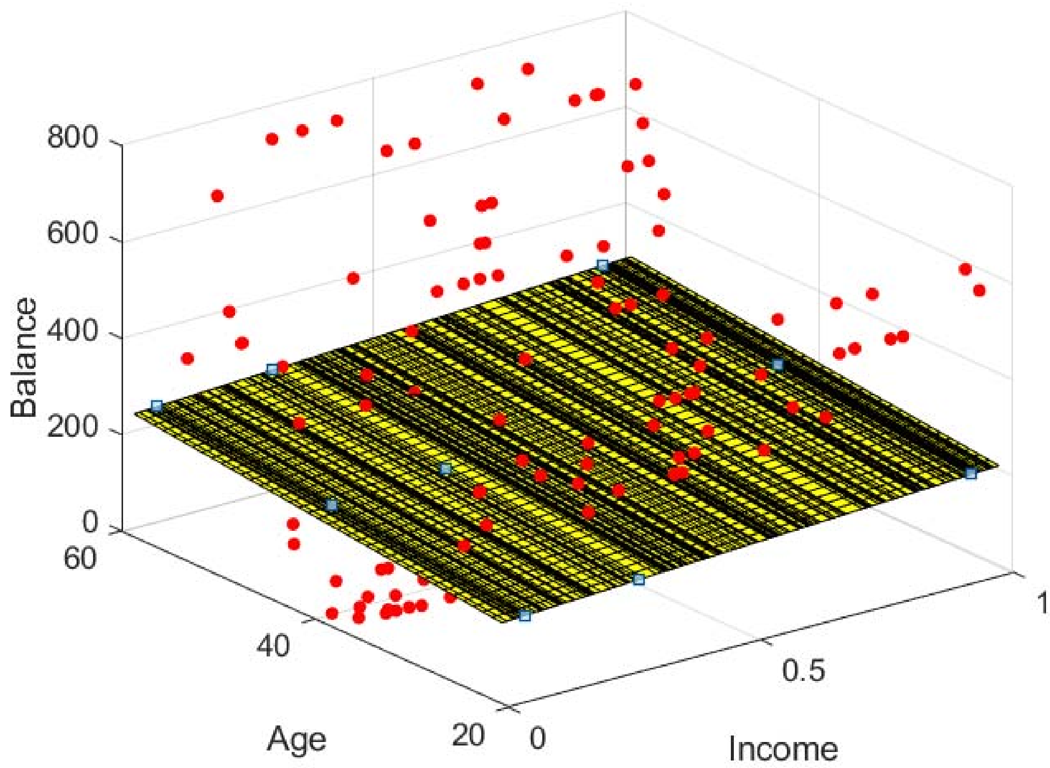
$$Min_{(u)} \quad J(\mathrm{u}) = \quad (Y - Au)^T \, (Y - Au)$$



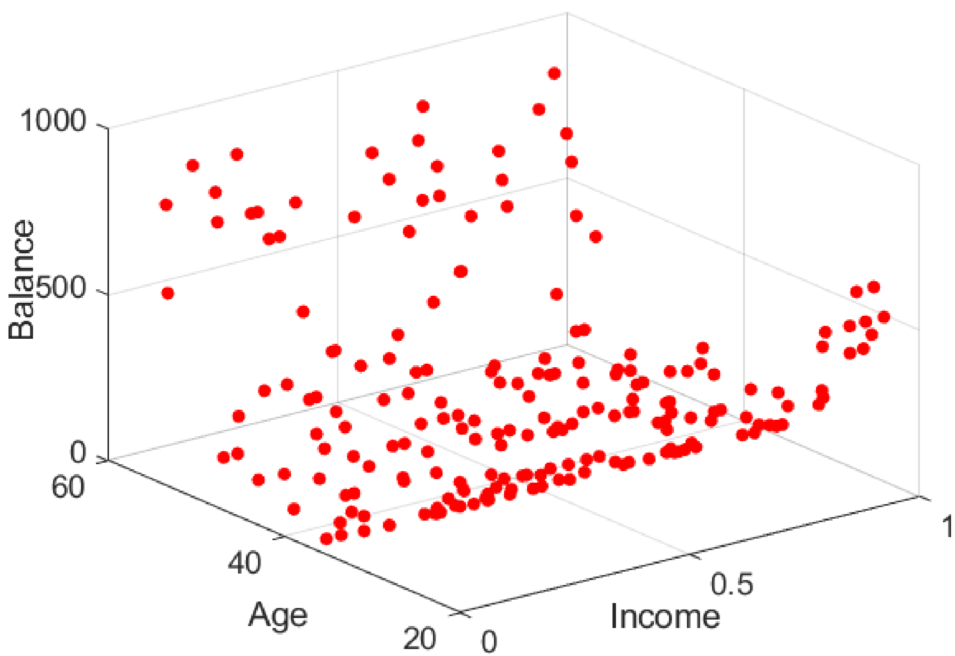
1000
500
0
Balance
60
40
20
Age
0
0.5
1
Income

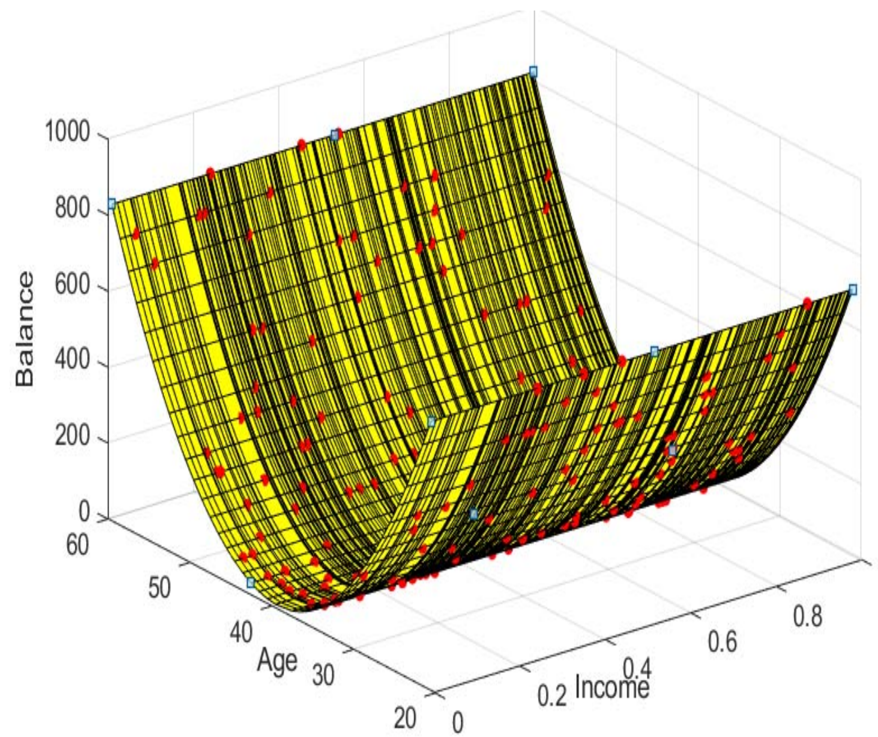
Quadratic Regression model with two variables





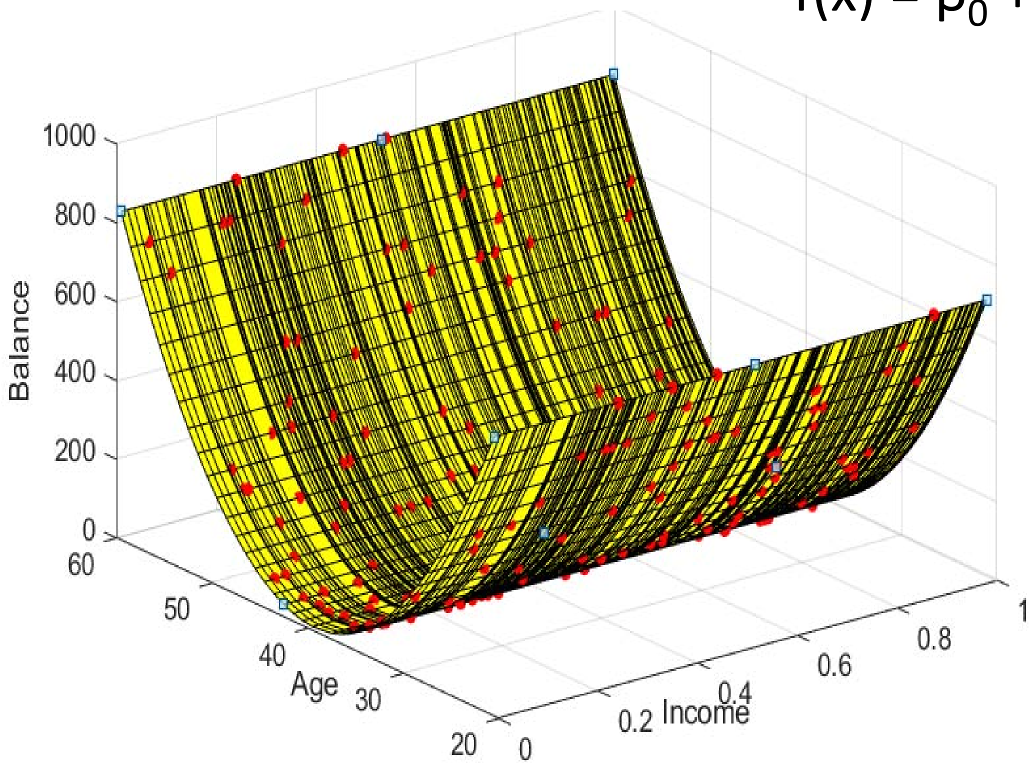
Quadratic Regression model with two variables





Quadratic Regression model with two variables

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_2^2 + \beta_5 x_1^2$$



For given Training Set T = { (x_{11}, x_{12}, ,y_1), (x_{21},x_{22}, y_2),….., (x_{n1},x_{n2}, y_n)}, we need to solve

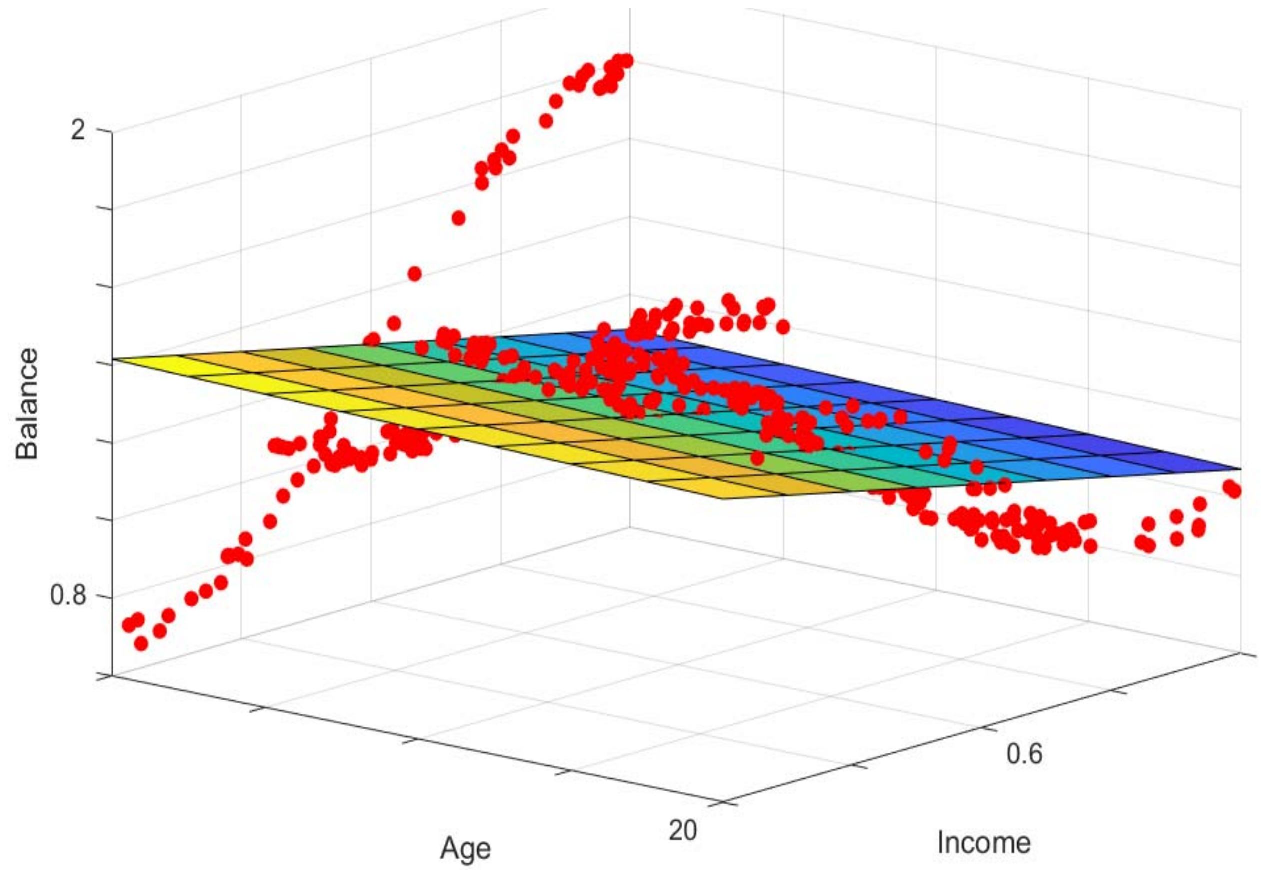
$$Min \quad J(\beta_5,\beta_4,\beta_3,\beta_2,\beta_1,\beta_0) = \quad \sum_{i=1}^{n} (yi-(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \beta_4 x_{2i}^2 + \beta_5 x_{1i}^2))^2 \qquad .....(1)$$

$$u= \begin{bmatrix} \beta_5 \\ \beta_4 \\ \beta_3 \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} \qquad A = \begin{bmatrix} {x_{11}}^2 & {x_{21}}^2 & x_{11} x_{12} & x_{12} & x_{11} & 1 \\ {x_{21}}^2 & {x_{22}}^2 & x_{21} x_{22} & x_{22} & x_{21} & 1 \\ & & & & & \\ {x_{n1}}^2 & {x_{n2}}^2 & x_{n1} x_{n2} & x_{n2} & x_{n1} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ y_n \end{bmatrix}$$

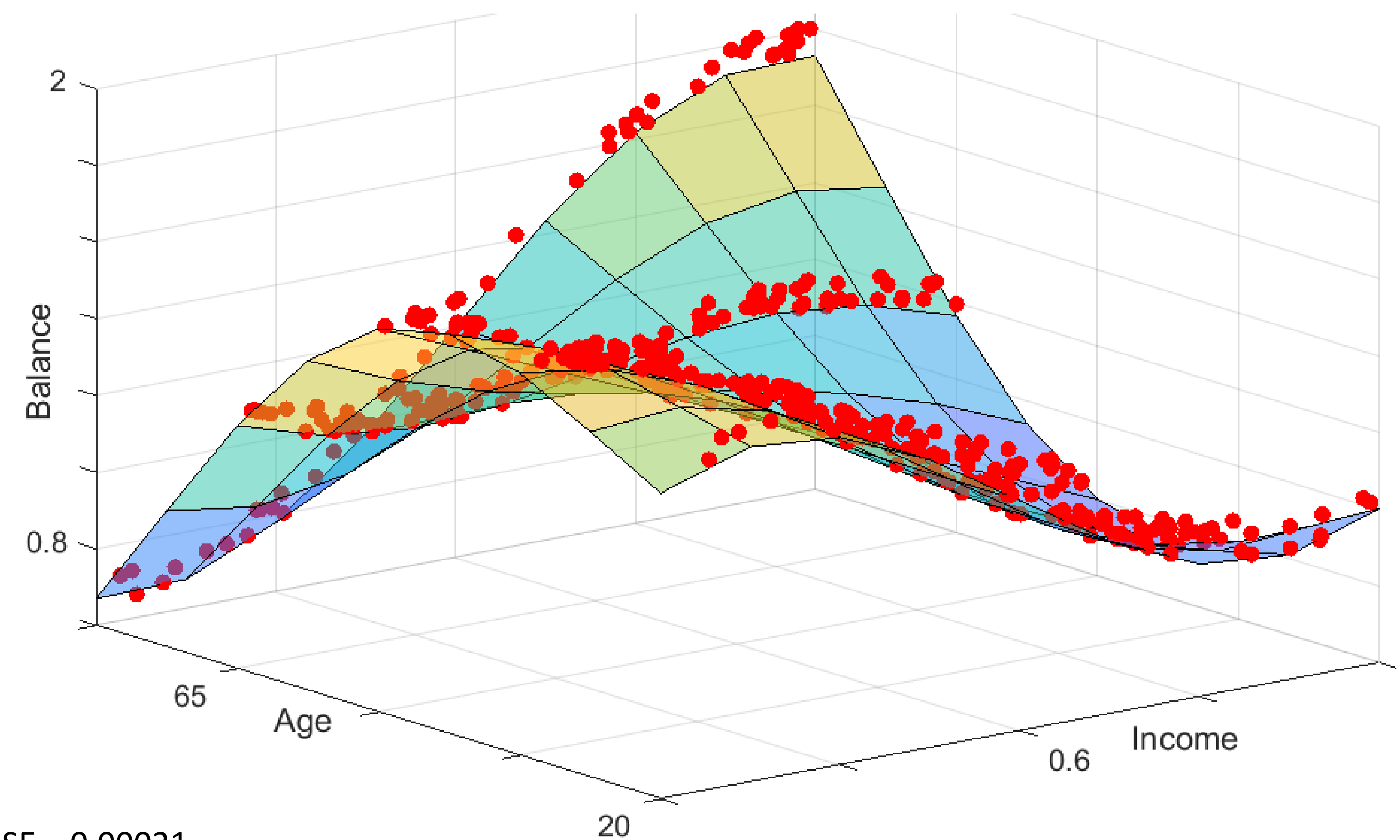
$$u= \begin{bmatrix} \beta_5 \\ \beta_4 \\ \beta_3 \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} = (A^T A)^{-1} \; A^T Y$$

The Least Square problem reduces to

$$Min_{(u)} \quad J(u) = \quad (Y - Au)^T (Y - Au)$$



Train RMSE = 0.4621



Train RMSE = 0.00021

Multiple Regression model working with k variables

x_1	x_2	x3	--	x_k	y Balance (thousand dollar)
32	0.550798	283	--	2	5.651202
22	0.708148	483	-	3	7.321263
45	0.290905	514	-	4	5.167304
78	0.510828	681	-	3	5.609367
54	0.892947	357	-	2	9.406379
39	0.896293	569	-	4	9.379439
42	0.125585	259	-	2	2.734997
51	0.207243	512	-	2	4.876649
21	0.051467	266	-	5	3.584138
19	0.44081	491	-	3	5.437239

Multiple Regression model working with k variables

Linear Function:-

$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \ldots + \beta_k x_k$

For given Training Set T = { $(x_{11}, x_{12}, \ldots, x_{1k}, y_1), (x_{21}, x_{22}, \ldots, x_{2k}, y_2), \ldots, (x_{n1}, x_{n2}, \ldots, xnk, y_n)$}, we solve

Min $J(\beta_k, \ldots, \beta_4, \beta_3, \beta_2, \beta_1, \beta_0) = \sum_{i=1}^{n} (yi - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \ldots + \beta_k x_k))^2$

$$u = \begin{bmatrix} \beta_k \\ \\ \\ \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} \qquad A = \begin{bmatrix} x_{11} & x_{21} & -- & x_{1k} & 1 \\ x_{21} & x_{22} & -- & x_{2k} & 1 \\ \\ \\ \\ x_{n1} & x_{n2} & -- & x_{nk} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem reduces to

$$Min_{(u)} \quad J(\mathrm{u}) = \quad (Y - Au)^T (Y - Au)$$

$$\boxed{u = \begin{bmatrix} \beta_k \\ \\ \\ \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} = (A^T A)^{-1} \; A^T \mathrm{Y}}$$

https://colab.research.google.com/drive/1APfTBXi3U_1ADX1PmLkJ8DyhVCQBegtE?usp=sharing