2-تقدير معالم نموذج الانحدار الخطي المتعدد:

تستخدم طريقة المربعات الصغرى لتقدير المعالم $B_0, B_1, ..., B_k$.

وتكون معادلة انحدار المربعات الصغرى كما يلي:

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ik} + e_i$$

حيث: $i = 1,2 \dots n$

$b_0, b_1, b_2, \dots b_k$ *هي القيم المقدرة للمعالم* $B_0, B_1, \dots, B_k$.

e_i تمثل البواقي.

يمكن كتابة متجهات ومصفوفات المعادلة السابقة كمايلي:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & . & x_{1k} \\ 1 & x_{21} & x_{22} & . & x_{2k} \\ 1 & x_{31} & x_{32} & . & x_{3k} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & x_{nk} \end{pmatrix} . \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ . \\ . \\ b_k \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ . \\ . \\ e_n \end{pmatrix}$$

وباستخدام رموز المصفوفات والمتجهات يمكن اختصار نموذج الانحدار السابق كما يلي: $y = xb + e$

وبالرجوع إلى طريقة المربعات الصغرى لتقدير معالم نموذج الانحدار الخطي المتعدد التي تجعل مجموع مربعات الأخطاء أقل ما يمكن.

$$\sum_{i=1}^{n} e_i^2 = \sum (y_i - \hat{y}_i)^2$$

ولدينا أيضا:

$$e^T e = (e_1 e_2 \dots \dots e_n) \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{pmatrix}$$

$$= e_1^2 + e_2^2 + \cdots + e_n^2$$

$$= \sum_{i=1}^{n} e_i^2$$

وبما أن : $y=xb+e$

ومنه $e=y-xb$

$$e^T e = (y - xb)^T \ (y-xb)$$

وبعد القيام بالعمليات الحسابية اللازمة نصل إلى:

$$e^T e = y^T y - 2b^T \,.\, x^T y + b^T \, x^T x \, b \, \dots\dots\dots (*)$$

وبتفاضل المعادلة (*) بالنسبة لـ b وجعل الناتج يساوي الصفر نجد:

$$\frac{\delta e^T e}{\delta b} = -2x^T y + 2x^T.xb = 0 \quad \Longrightarrow \quad x^T xb = x^T.y$$

وبضرب طرفي هذه المعادلة بـ $(x^T x)^{-1}$ نحصل على:

$$(x^T x)^{-1}.x^T xb = (x^T x)^{-1}.\, x^T y$$

ومنه يكون:

$$\hat{b} = (x^T x)^{-1} \cdot x^T y$$

مثال01:

تم التعاقد مع شركة تسويق لتقدير مقدار إنفاق الأسرة على الغذاء بالاعتماد على دخل وحجم الأسرة، وتشمل البيانات التالية الإنفاق الشهري على الطعام y (بآلاف الدولارات) في مقابل الدخل الشهري $x1$ (بآلاف الدولارات) وحجم العائلة $x2$ وذلك لعدد 15 عائلة مختارة عشوائيا من منطقة معينة.

إنفاق الأسرة (y)	الدخل الشهري ($x1$)	حجم العائلة ($x2$)
0.43	2.1	3
0.31	1.1	4
0.32	0.9	5
0.46	1.6	4
1.25	6.2	4
0.44	2.3	3
0.52	1.8	6
0.29	1	5
1.29	1	5

2	2.4	0.35
4	1.2	0.35
3	4.7	0.78
2	3.5	0.43
3	2.9	0.47
4	*1.4*	*0.38*

المطلوب:

-قدر معادلة الانحدار المتعدد؟

الحل:

-تقدير معادلة الانحدار المتعدد:

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2}$$

نطبق العلاقة التالية: $\hat{b} = (x^T x)^{-1} \cdot x^T y$

إيجاد المصفوفة:

$$(x^T x)^{-1} = \begin{pmatrix} 1.360838 & -0.092616 & -0.282230 \\ -0.092616 & 0.01656 & 0.012612 \\ -0.282230 & 0.012612 & 0.067341 \end{pmatrix}$$

ثم إيجاد $(x^T y)$:

$$(x^T y) = \begin{pmatrix} 8.07 \\ 32.063 \\ 28.96 \end{pmatrix}$$

نقوم بتطبيق العلاقة: $\hat{b}=(x^T x)^{-1}. x^T y$

لنحصل على:

$$=\begin{pmatrix} 1.360838 & -0.092616 & -0.282230 \\ -0.092616 & 0.01656 & 0.012612 \\ -0.282230 & 0.012612 & 0.067341 \end{pmatrix}\begin{pmatrix} 8.07 \\ 32.063 \\ 28.96 \end{pmatrix}$$

$$\hat{b}=\begin{pmatrix} -0.16045 \\ 0.14872 \\ 0.07691 \end{pmatrix}$$

ومنه فإن معادلة الانحدار هي:

$$\hat{y}_i = -0.16045 + 0.14872x_1 + 0.07691x_2$$

- **تفسير معاملات نموذج الانحدار الخطي المتعدد:**

يمكننا تفسير المعامل الثابت (b_0) على أنه يمثل القيمة المقدرة للمتغير التابع عندما تكون قيم المتغيرات المستقلة مساوية للصفر أي ($x_1, x_2, \ldots\ldots,\ldots x_k$ معدومة).

بالنسبة للمعامل b_1 : يمثل مقدار التغير في القيمة المتوقعة للمتغير التابع $\hat{y}_i$ لما يتغير x_1 بوحدة واحدة مع افتراض ثبات المتغيرات المستقلة الأخرى ($x_2, x_3,\ldots,x_k$).

بالنسبة للمعامل b_2 : يمثل مقدار التغير في القيمة المتوقعة للمتغير التابع $\hat{y}_i$ لما يتغير x_2 بوحدة واحدة مع افتراض ثبات المتغيرات المستقلة الأخرى ($x_1, x_3,\ldots,x_k$).

وهكذا يمكن تفسير باقي معاملات الانحدار الخطي المتعدد.

- **خصائص مقدرات المربعات الصغرى:**

تتميز مقدرات المربعات الصغرى بالخصائص التالية:

أ-خاصية عدم التحيز:

وتعني أن القيمة المتوقعة لكل عنصر من عناصر المتجه(b) تساوي العنصر المقابل في متجه المعالم الحقيقية (B) أي أن:

$$E(b)= B$$

ب-خاصية الخطية:

بما أن$(x^T x)^{-1}. x^T y$ هي مصفوفة أرقام ثابتة فان(b) دالة خطية لـ (Y) وعليه فان مقدرات المربعات الصغرى هي مقدرات خطية.

ج-خاصية الكفاءة:

إن تباين مقدرات المربعات الصغرى اقل من أو يساوي تباين أية مقدرات أخرى خطية وغير متحيزة. وتباين مقدرات المربعات الصغرى يعطى بالعلاقة التالية:

$$\sigma^2(b)= \sigma^2 (x^T x)^{-1}$$

- **تباين مقدارت المربعات الصغرى:**

يعطي تباين مقدرا ت المربعات الصغرى بالصيغة التالية : $\sigma^2(b)= S_e^2 (x^T x)^{-1}$

حيث أن S_e^2 هو مقدر غير متحيز لـ σ^2

$$S_e^2 = \frac{SSE}{n-k-1} = \frac{y^T y - b^T(x^T y)}{n-(k+1)}$$