The Multivariate Regression

الانحدار متعدد المتغيرات

نعلم أن نموذج الانحدار الخطي المتعدد للمتغير Y (متغير الاستجابة) على مجموعة المتغيرات التوضيحية $(X_1, X_2 \dots X_p)$ له الشكل الخطي الآتي:

$$Y_j = B_0 + B_1 X_{1j} + B_2 X_{2j} + \dots + B_p X_{pj} + U_j$$

$$j = 1, 2, \dots, n$$

$$\dots\dots\dots (1)$$

نلاحظ في المعادلة أعلاه أن المشاهدة j للمتغير الاستجابة Y هي عبارة عن دالة خطية للمشاهدة j للمتغيرات التوضيحية $X_1, X_2 \dots X_n$ وأخطاء العشوائي U_j. إذ أن U_j يتبع التوزيع الطبيعي الاعتيادي المتغير عيوسط = 0 وتباين σ^2 أي أن

$$U_j \sim N(0, \sigma^2)$$

سنوضح تعميم المعادلة (1) أعلاه إلى نموذج آخر يسمى بنموذج الانحدار المتعدد المتغيرات الخطي والذي له الشكل الآتي:

$$Y_\alpha = B_{0\alpha} + B_{1\alpha} X_{1\alpha} + B_{2\alpha} X_{2\alpha} + \dots + B_k X_{k\alpha} + U_\alpha$$

$$\alpha = 1, 2, \dots, p$$

$$\dots\dots\dots (2)$$

إذ أن U_α يتبع متجه أخطاء العشوائية (Y_α) ومتجه متغير الاستجابة $(p \times 1)$

Scanned with CamScanner

عندنوز (n) ن مشاهدات متغير الاستجابة والمتغيرات التوضيحية
يمكن كتابة النموذج (1) بشكل مصفوفات وكالاتي:

$$Y = XB + U$$

اذ ان المصفوفة Y تمثل مصفوفة مشاهدات المتغيرات المعتمدة:

$$\underline{Y}_1, \underline{Y}_2 \dots, \underline{Y}_n$$

$(p \times n)$ من ذات المصفوفة مشاهدات المتغيرات المعتمدة

هذا يعني ان مصفوفة مشاهدات المتغيرات المعتمدة

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1n} \\ y_{21} & y_{22} & \dots & y_{2n} \\ \vdots & & & \\ y_{p1} & y_{p2} & \dots & y_{pn} \end{bmatrix}_{p \times n} = (\underline{Y}_1 \ \underline{Y}_2 \dots \underline{Y}_n)$$

اي ان:

اما مصفوفة المشاهدات للمتغيرات التوضيحية

$$X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & & & \\ x_{k1} & x_{k2} & \dots & x_{kn} \end{bmatrix}_{(k+1) \times n}$$

وتمثل مصفوفة B مصفوفة معلمات اذا متغيرات الاستجابة ← اضافة الى عمود معلمة المقطع $X_1, X_2 \dots X_n$ وهي مصفوفة ذات سعة $Y_1, Y_2 \dots Y_n$ لكل المتغيرات التوضيحية وناتجة الشكل الاتي $P \times (k+1)$

$$B = \begin{bmatrix} B_{01} & B_{11} & B_{21} & \dots & B_{k1} \\ B_{02} & B_{12} & B_{22} & & B_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ B_{0p} & B_{1p} & B_{2p} & & B_{kp} \end{bmatrix}_{P \times (k+1)} = \begin{bmatrix} B_0 & B_1 \\ p \times 1 & p \times k \end{bmatrix}$$

Scanned with CamScanner

- اذ ان المتجه $\underline{\beta_0}$ هو متجه ذات سعة $(p \times 1)$ والذي يمثل المعدد الاول ن المصفوفة β كل عنصر منه يمثل معاملات القطع لكل معادلة

- اما المصفوفة β_1 فهي مصفوفة ذات سعة $p \times k$ والتي تمثل مصفوفة معاملات الانحدار لكل معادلة

اذ ان الصف الاول ن المصفوفة β_1 يمثل معاملات انحدار متغير الاستجابة (Y_1) عن المتغيرات التوضيحية $(X_1, X_2 \dots X_k)$

اما الصف الثاني ن المصفوفة β_1 يمثل معاملات انحدار متغير الاستجابة (Y_2) عن المتغيرات التوضيحية $(X_1, X_2 \dots X_k)$ وهكذا ...

- اما مصفوفة U فهي تمثل مصفوفة الاخطاء العشوائية ذات سعة $(p \times n)$ كل عمود منها يمثل المتجه العشوائي المتعدد المتغيرات بـ p ن المتغيرات يمثل صفوف ومصفوفة تباين $\neq$ والمصفوفة $\neq$ ذات سعة $p \times p$ وهي مصفوفة متماثلة واكيدة الايجابية حيث ان n اكبر من $(p+k+1)$

$$U_{p \times n} = (\underline{U_1} \ \underline{U_2} \dots \underline{U_n})$$

$$Var(\underline{U_1}) = \hat{\sigma} I = Var(\underline{U_2}) = \hat{\sigma} I = \dots = Var(\underline{U_n})$$

$$E(\underline{U_1}) = 0 \ , \ E(\underline{U_2}) = 0 \dots E(\underline{U_n}) = 0$$

Scanned with CamScanner

Parameter estimation of Multivariate Linear Regression :-

تقدير معلمات نموذج الانحدار الخطي المتعدد

قبل البدء لتقدير المعلمات عكتب كتابة النموذج ( ) المعادلة بالشكل الاتي:-

$$\underline{Y_i} = B\,\underline{x_i} + \underline{u_i} \quad \text{for } i = 1, \dots, n \qquad \text{---} (3)$$

$$\underline{Y_i} = \begin{bmatrix} Y_{1i} \\ Y_{2i} \\ \vdots \\ Y_{pi} \end{bmatrix} \; ; \; \underline{x_i} = \begin{bmatrix} 1 \\ x_{1i} \\ x_{2i} \\ \vdots \\ x_{pi} \end{bmatrix} \; ; \; \underline{u_i} = \begin{bmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{pi} \end{bmatrix}$$

① Least Squares Method

طريقة المربعات الصغرى

تعتمد طريقة المربعات الصغرى على مبدأ هو جعل مربعات البواقي اقل ما يمكن

فاذا افترضنا ان نموذج الانحدار متعدد المتغيرات المقدر للمجتمع هو:

$$\underline{\hat{Y}_i} = \hat{B}\,\underline{x_i}$$

وعليه تكون متجه البواقي بالشكل الاتي

$$\underline{e_i} = \underline{Y_i} - \underline{\hat{Y}_i} = \underline{Y_i} - \hat{B}\,\underline{x_i}$$

وعليه تكون دالة الهدف في انحدار متعدد المتغيرات بطريقة المربعات الصغرى كالاتي :

$$\min Q = \sum_{i=1}^{n} \underline{e_i}'\,\underline{e_i}$$

$$\text{or } \min Q = \sum_{i=1}^{n} tr\, \underline{e_i}\,\underline{e_i}'$$

Scanned with CamScanner

بعد اشتقاق دالة الخطأ حسب المصفوفة المعاملات $(\hat{B})$ وتساوي المشتقين الأول والثاني بصفر نجد $\hat{B}$ التي تجعل Q دالة الخطأ أقل ما يمكن وتكون بالشكل الآتي:

$$\hat{B} = \sum_{i=1}^{n} Y_\alpha X_\alpha \left(\sum_{i=1}^{n} x_\alpha x_\alpha\right)^{-1} = Y X' (X X')^{-1}$$

$$\therefore \hat{B} = Y_{n \times p} X' (X X')^{-1}$$

where:

$$\sum_{i=1}^{n} Y_i \underline{X}'_i = Y X' = \begin{bmatrix} \Sigma Y_{1i} & \Sigma Y_{1i} X_{1i} & \cdots & \Sigma X_{1i} X_{k,i} \\ \Sigma Y_{2i} & \Sigma Y_{2i} X_{1i} & \cdots & \Sigma Y_{2i} X_{ki} \\ \vdots & & & \\ \Sigma Y_{pi} & \Sigma Y_{pi} X_{1i} & \cdots & \Sigma Y_{pi} X_{k,i} \end{bmatrix}$$

$P_{*(k+1)}$

$$\Sigma \underline{X}_i \underline{X}'_i = X X' = \begin{bmatrix} n & \Sigma X_{1i} & \Sigma X_{2i} & \cdots & \Sigma X_{ki} \\ & \Sigma X_{1i}^2 & \Sigma X_{1i} X_{2i} & \cdots & \Sigma X_{1i} X_{ki} \\ & & \ddots & & \\ & & & & \Sigma X_{ki}^2 \end{bmatrix}$$

اما مصفوفة معاملات الانحدار المقدرة فتكون بالشكل الآتي:

$$\hat{B} = \begin{bmatrix} \hat{B}_{01} & \hat{B}_{11} & \hat{B}_{21} & \cdots & \hat{B}_{k1} \\ \hat{B}_{02} & \hat{B}_{12} & \hat{B}_{22} & \cdots & \hat{B}_{k2} \\ \vdots & & & & \\ \hat{B}_{0p} & \hat{B}_{1p} & \hat{B}_{2p} & \cdots & \hat{B}_{kp} \end{bmatrix}$$

Scanned with CamScanner

اذ ان الصف الاول بالصفوف ($\hat{\beta}$) تمثل معاملات الانحدار المقدرة لانحدار متغير الاستجابة (Y_1) على المتغيرات التوضيحية ($X_1, X_2 \dots X_k$)
والصف الثاني تمثل معاملات الانحدار المقدرة لانحدار متغير الاستجابة (Y_2) على جميع المتغيرات التوضيحية ($X_1 X_2 \dots X_k$) وهكذا بالنسبة لبقية الصفوف

اما متجه متغيرات الاستجابة المقدرة ($\hat{Y}_i$) فيكون

$$\hat{Y}_i = \hat{\beta} \underline{x}_i$$

Scanned with CamScanner