Αν το μοντέλο είναι $\underline{y} = \mathrm{X}\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim N_n(0, \sigma^2 I)$, $\underline{\beta} \in \Re^p$, rank$(\mathrm{X}) = p < n$,

έχω ότι $\hat{\underline{\beta}} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{y}$, $\hat{\sigma}^2 = \dfrac{\underline{y}^T(I - \mathrm{P}_V)\underline{y}}{n-p}$,

$\underline{\mu} = \mathrm{X}\underline{\beta}$ □ $\underline{\beta} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{\mu}$ □ Η προβολή του $\underline{\mu}$ στο $\Re^p$.

Ισχύει το παρακάτω θεώρημα ανάλογα με το μοντέλο #1:

<u>Θεώρημα Gauss-Markov:</u>
Έστω $\underline{d} \in \Re^p$, σταθερό. Τότε η εκτιμήτρια $\underline{d}^T\hat{\underline{\beta}}$ είναι **<u>η</u>** μοναδική γραμμική εκτιμήτρια του $\underline{d}^T\underline{\beta}$ με ελάχιστη διασπορά.
<u>Απόδειξη:</u> Αφού $\underline{\beta} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{\mu}$ □ $\underline{d}^T\underline{\beta} = \underbrace{\underline{d}^T(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T}_{c}\underline{\mu} = \underline{c}^T\underline{\mu}$ και άρα εφαρμόζω το θεώρημα Gauss-Markov για το μοντέλο #1.

<u>Πόρισμα:</u>
Η εκτιμήτρια ελαχίστων τετραγώνων $\hat{\underline{\beta}}$ είναι Α.Ο.Ε.Δ. του $\underline{\beta}$ κατά συνιστώσα.

<u>Εφαρμογές: #1. Απλό γραμμικό μοντέλο</u>
Έχω ότι $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \ldots, n$.
Γνωρίζω ότι $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$, $\hat{\beta}_1 = \dfrac{\sum(x_i - \bar{x})y_i}{\sum(x_i - \bar{x})^2}$.

Θεωρώ $\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $\mathrm{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ και $\underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$, οπότε $\underline{y} = \mathrm{X}\underline{\beta} + \underline{\varepsilon}$, $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$.

Έχω λοιπόν ότι $\hat{\underline{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{y}$.

Όμως, $\mathrm{X}^T\mathrm{X} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \end{pmatrix}\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}$.

Η ορίζουσα του $\mathrm{X}^T\mathrm{X}$ είναι $n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2 = n\sum_{i=1}^n x_i^2 - n^2\bar{x}^2 = n\sum_{i=1}^n (x_i - \bar{x})^2$.

Οπότε $(\mathrm{X}^T\mathrm{X})^{-1} = \dfrac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2}\begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix} = \dfrac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2}\begin{bmatrix} \sum_{i=1}^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}$

Άρα $(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T = = \dfrac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2}\begin{bmatrix} \sum_{i=1}^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}\begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \end{pmatrix}$

$$= \frac{1}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}\begin{bmatrix} \sum x_i^2 - n\bar{x}x_1 & \sum x_i^2 - n\bar{x}x_2 & \dots & \sum x_i^2 - n\bar{x}x_n \\ -n\bar{x}+nx_1 & -n\bar{x}+nx_2 & \dots & -n\bar{x}+nx_n \end{bmatrix}.$$

Άρα, $\underline{\hat{\beta}} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{y} = \dfrac{1}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}\begin{bmatrix} \sum x_i^2 - n\bar{x}x_1 & \dots & \sum x_i^2 - n\bar{x}x_n \\ -n\bar{x}+nx_1 & \dots & -n\bar{x}+nx_n \end{bmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$= \frac{1}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}\begin{bmatrix} \left(\sum x_i^2 - n\bar{x}x_1\right)y_1 + \dots + \left(\sum x_i^2 - n\bar{x}x_n\right)y_n \\ (-n\bar{x}+nx_1)y_1 + \dots + (-n\bar{x}+nx_n)y_n \end{bmatrix}$$

$$= \frac{1}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}\begin{bmatrix} n\bar{y}\sum x_i^2 - n\bar{x}\sum x_i y_i \\ -n^2\overline{xy} + n\sum x_i y_i \end{bmatrix}.$$

Άρα, για παράδειγμα $\hat{\beta}_1 = \dfrac{n\sum x_i y_i - n^2\overline{xy}}{n\sum (x_i-\bar{x})^2} = \dfrac{\sum x_i y_i - n\overline{xy}}{\sum (x_i-\bar{x})^2} = \dfrac{\sum (x_i-\bar{x})y_i}{\sum (x_i-\bar{x})^2}$.

<u>#2. Πολλαπλή Γραμμικλη Παλινδρόμηση</u>

$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i$, $i = 1, \dots, n$.

Γνωρίζουμε ότι $SS_{tot} = SS_{reg} + SS_{res}$, όπου

$$SS_{tot} = \sum_{i=1}^{n}(y_i - \bar{y})^2 ,$$

$$SS_{reg} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2 ,$$

$$SS_{res} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 .$$

Θα αποδείξουμε αυτήν την ανάλυση με προβολές.

Μπορώ να γράψω το μοντέλο

$\underline{y} = \mathrm{X}\underline{\beta} + \underline{\varepsilon} = \underline{\mu} + \underline{\varepsilon}$, $\underline{\mu} \in V$ =χώρος που παράγεται από τις στήλες του X .

Έχω ότι το εκτιμώμενο μοντέλο είναι $\underline{\hat{y}} = \underline{\hat{\mu}} = \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix}$, οπότε

$SS_{tot} = \|\underline{y} - \bar{y}\cdot\underline{1}\|^2$, $SS_{reg} = \|\underline{\hat{y}} - \bar{y}\cdot\underline{1}\|^2$, $SS_{res} = \|\underline{y} - \underline{\hat{y}}\|^2$.

Αφού V γνήσιος υπόχωρος του $\Re^n$, έχω ότι $\Re^n = V \oplus V^{\perp}$.

Αλλά αν $M = <\underline{1}>$, τότε $M \subset V$. Επομένως $V = M \oplus V/M$.

Άρα $\Re^n = M \oplus V/M \oplus V^{\perp}$.

Αν $\underline{y} \in \Re^n$ □ $\underline{y} = \mathrm{P}_M\underline{y} + \mathrm{P}_{V/M}\underline{y} + \mathrm{P}_{V^{\perp}}\underline{y}$ και οι προβολές είναι ανά δύο ορθογώνιες.

Άρα $\|\underline{y}\|^2 = \|\mathrm{P}_M\underline{y}\|^2 + \|\mathrm{P}_{V/M}\underline{y}\|^2 + \|\mathrm{P}_{V^{\perp}}\underline{y}\|^2$ (□).

Έχω ότι $\mathrm{P}_M\underline{y} = \dfrac{(\underline{1}^T\cdot\underline{y})}{\underline{1}^T\cdot\underline{1}}\cdot\underline{1} = \bar{y}\cdot\underline{1}$ □ $\|\mathrm{P}_M\underline{y}\|^2 = n\bar{y}^2$.

$$\mathrm{P}_{V/M}\,\underline{y} = \mathrm{P}_V\,\underline{y} - \mathrm{P}_M\,\underline{y} \;\square\; \|\mathrm{P}_{V/M}\,\underline{y}\|^2 = \|\mathrm{P}_V\,\underline{y} - \mathrm{P}_M\,\underline{y}\|^2 = \|\hat{\underline{y}} - \overline{y}\cdot\underline{1}\|^2$$

$$\|\mathrm{P}_{V^{\perp}}\,\underline{y}\|^2 = \|\underline{y} - \mathrm{P}_V\,\underline{y}\|^2 = \|\underline{y} - \hat{\underline{\mu}}\|^2 = \|\underline{y} - \hat{\underline{y}}\|^2 .$$

Άρα, ($\square$) $\square$ $\|\underline{y}\|^2 = n\overline{y}^2 + \|\hat{\underline{y}} - \overline{y}\cdot\underline{1}\|^2 + \|\underline{y} - \hat{\underline{y}}\|^2$

$$\square\; \sum_{i=1}^{n} {y_i}^2 - n\overline{y}^2 = SS_{reg} + SS_{res}$$

$$\square\; SS_{tot} = SS_{with} + SS_{res} .$$