Ασκήσεις 2 (Πολλαπλή παλινδρόμηση)

Ασκ. 1: Επαληθεύστε τη διασπορά του σφάλματος πρόβλεψης $\tilde{\varepsilon}_{n+1} = y_{n+1} - \tilde{y}_{n+1}$ της απλής παλινδρόμησης, μέσω της αντίστοιχης διασποράς της πολλαπλής παλινδρόμησης.

Ασκ. 2: Έστω $X_{(n,p)}$ ένας πίνακας τάξης p και $\hat{Y}$ η ορθογώνια προβολή ενός διανύσματος Y του $\mathbb{R}^n$ στον υπόχωρο που παράγεται από τις στήλες του X. Συμβολίζουμε με 1_n το διάνυσμα του $\mathbb{R}^n$ που αποτελείται από μονάδες.

α) Εκφράστε το εσωτερικό γινόμενο $\langle Y, 1_n \rangle$ ως συνάρτηση των y_i.

β) Έστω $\hat{\varepsilon} = Y - \hat{Y}$ και υποθέτουμε ότι ένα από τα διανύσματα στήλες του X είναι της μορφής $(c, c, \ldots, c)^t$ με $c \neq 0$.
Βρείτε το $\langle \hat{\varepsilon}, 1_n \rangle$.

γ) Συμπεράνετε ότι αν η σταθερά b_1 συμπεριλαμβάνεται στο μοντέλο της γραμμικής παλινδρόμησης τότε $\bar{y} = \overline{\hat{y}}$ (δειγμ. μέσοι)

Ασκ 3.: Εξετάζουμε την μεταβλητή y ως συνάρτηση 2 ανεξάρτ. μεταβλητών x και z. Διαθέτουμε n-παρατηρήσεις των μεταβλητών αυτών.
Συμβολίζουμε $X = [1_n \ \underline{x} \ \underline{z}]$, όπου $\underline{x} = (x_1, \ldots, x_n)^t$, $\underline{z} = (z_1, \ldots, z_n)^t$.

α) Πήραμε τα παρακάτω αποτελέσματα:

$$X^t X = \begin{bmatrix} 25 & 0 & 0 \\ ? & 9.3 & 5.4 \\ ? & ? & 12.7 \end{bmatrix}, \quad (X^t X)^{-1} = \begin{bmatrix} 0.04 & 0 & 0 \\ 0 & 0.1428 & -0.0607 \\ 0 & -0.0607 & 0.1046 \end{bmatrix}$$

(i) Βρείτε τις τιμές που λείπουν.

(ii) Ποια είναι η τιμή του n ?

(iii) Βρείτε τον εμπειρικό συντελεστή γραμμικής συσχέτισης $r_{x,z}$.

β) Η γραμμική παλινδρόμηση της Y στο $(1_n, \underline{X}, \underline{Z})$ έδωσε

$$Y = -1.6\, 1_n + 0.61\, \underline{X} + 0.46\, \underline{Z} + \hat{\varepsilon},$$

όπου $SSE = 0.3$.

(i) Βρείτε τον εμπειρικό μέσο $\bar{Y}$.

(ii) Υπολογίστε τα SSR, SST, R^2 και $\overline{R^2}$.

Ασκ. 4.

Θεωρούμε το μοντέλο γραμμικής παλινδρόμησης $Y = X\beta + \varepsilon$, όπου $Y \in \mathbb{R}^n$, ο X είναι $n \times p$ και αποτελείται από p ορθογώνια διανύσματα, $\beta \in \mathbb{R}^p$ και $\varepsilon \in \mathbb{R}^n$. Έστω $X = [\underset{n\times q}{Z} \,\vdots\, \underset{n\times(p-q)}{U}]$

Πήραμε μέσω της μεθόδου των συνήθων ελαχίστων τετραγώνων τα εξής:

$$\hat{Y}_X = \hat{\beta}_1^X X_1 + \dots + \hat{\beta}_p^X X_p$$

$$\hat{Y}_Z = \hat{\beta}_1^Z X_1 + \dots + \hat{\beta}_q^Z X_q$$

$$\hat{Y}_U = \hat{\beta}_{q+1}^U X_{q+1} + \dots + \hat{\beta}_p^U X_p.$$

Συμβολίζουμε επίσης $SSR(A) = \|P_A Y\|^2$, όπου P_A ο πίνακας ορθογώνιας προβολής στον A.

α) Δείξτε ότι $SSR(X) = SSR(Z) + SSR(U)$

β) Δώστε μία έκφραση του $\hat{\beta}_1^X$ ως συνάρτηση του Y, του X_1 και $\|X_1\|$.

γ) Συμπεράνετε ότι $\hat{\beta}_1^X = \hat{\beta}_1^Z$