«ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΟΣ: ΠΑΡΑΤΗΡΗΣΕΙΣ»

Τοῦ κ. ΕΠΑΜΕΙΝΩΝΔΑ Ε. ΠΑΝΑ (M. Sc.)

Βοηθοῦ τῆς Τακτικῆς ἕδρας τῆς Οἰκονομετρίας - Α.Σ.Ο.Ε.Ε.

1. Εἰσαγωγὴ

Ἀφορμὴ τοῦ παρόντος σημειώματος εἶναι τὸ ἄρθρον τοῦ Μ. Σακέλλη (1) δημοσιευθὲν εἰς τὰς Σπουδὰς καὶ ἡ ἀρχὴ τοῦ S. Johnson (1709 - 1784), κατὰ τὴν ὁποίαν «Ἔχομεν δύο εἰδῶν γνώσεις ἑνὸς ἀντικειμένου· ἡ μία εἶναι ὅτι γνωρίζομεν ἐμεῖς οἱ ἴδιοι τὸ ἀντικείμενον (δική μας πνευματικὴ ἀνακάλυψις) ἢ γνωρίζομεν ποῦ δυνάμεθα νὰ εὕρωμεν τὰ στοιχεῖα τοῦ ἀντικειμένου ποὺ διαφερόμεθα.

Τὸ πρῶτον κοινὸν ἐρώτημα ὅταν ἔρχεται — πιστεύω — κανεὶς εἰς ἐπαφὴν μὲ τὸ ἄρθρον τοῦ Σακέλλη εἶναι ποῖος εἶναι ὁ σκοπὸς τοῦ ἄρθρου αὐτοῦ;

Δύο ἀπαντήσεις ὑπάρχουν ἢ ὅτι ὁ συγγραφεὺς προσέθεσε κάτι τὸ νέον εἰς τὸ πρόβλημα τῆς πολυσυγγραμμικότητος, ἢ ὅτι παρουσιάζει μίαν γενικὴν περίληψιν τοῦ θέματος.

Ἀμφότεραι αἱ ἀπαντήσεις δὲν ἱκανοποιοῦν τὴν ἀρχὴν τοῦ Johnson καὶ δὲν τὴν ἱκανοποιοῦν διότι οὔτε κάτι νέον προσέθεσεν ἀλλὰ οὔτε καὶ μίαν γενικὴν περίληψιν τοῦ θέματος παρουσίασε καθὼς θὰ ἴδωμεν περαιτέρω.

Δυστυχῶς εἴμεθα ὑποχρεωμένοι νὰ ἐπισημάνωμεν τὸ γεγονὸς ὅτι σχεδὸν τὸ ὅλον ἄρθρον τοῦ Σακέλλη εἶναι ἀντιγραφὴ ἐργασίας τοῦ Giles (2) *. Σημειωτέον ὅτι ἐκ τοῦ συνόλου τῶν 384 στοίχων τοῦ περὶ οὗ ἄρθρου (1) μόνον οἱ 88 δὲν ἀνήκουν εἰς τὸν Giles, ἐνῶ τὸ μεγαλύτερον μέρος τῶν ὑπολοίπων στοίχων ἀναφέρεται εἰς τοὺς Farrar καὶ Clauber (3),

Παρὰ ταῦτα τὸ μέρος τὸ ληφθὲν ἀπὸ τὴν ἐργασίαν τοῦ Giles δὲν ἀπεδόθει πιστά, διότι ἀφ' ἑνὸς μὲν ὑπάρχουν μεταφραστικὰ λάθη, ὀφειλόμενα προφανῶς εἰς τὴν μὴ ἐπαρκὴν γνῶσιν τῆς ὀρολογίας τῆς γραμμικῆς ἀλγέβρας ὑπὸ τοῦ συγγραφέως τοῦ ἄρθρου, ἀφ' ἑτέρου λόγῳ ἐνδιαμέσων παραλείψεων.

---

* Σημ. Συντάξεως. Ὁ κ. Σακέλλης μᾶς ἀνεκοίνωσε δι' ἐπιστολῆς του (21) 4/77) ὅτι τὸ δημοσιευθὲν ἄρθρον του «Στ.πουδαί». 4ον, 1976) στηρίζεται εἰς τὸ ἔργον τοῦ D E..A. Gilés «Essays on Econometric Topics : From theory to practice».

Ὁ σκοπὸς τοῦ παρόντος συντόμου σημειώματος εἶναι ἀφ' ἑνὸς μὲν νὰ παρουσιάσῃ ἐν συντομίᾳ τὶς τελευταῖες μεθόδους ἐπάνω εἰς τὸ πρόβλημα τῆς πολυσυγγραμμικότητος καὶ ἀφ' ἑτέρου νὰ παρουσιάσῃ τὶς κριτικὲς ποὺ ἔχουν γραφῆ ἐπάνω εἰς τὸ θέμα, τὸ ὁποῖον παρουσίασεν ὁ Σακέλλης.

Ὑπόδειγμα : Ἔχομεν τὸ ἀκόλουθον γραμμικὸν ὑπόδειγμα :

$$y = x\beta + \varepsilon \qquad (1)$$

ὅπου :

y : εἶναι ἓν $T \times 1$ διαστάσεως διάνυσμα τῆς ἐξηρτημένης μεταβλητῆς
x : εἶναι εἷς $T \times N$ διαστάσεως μὴ στοχαστικὸς πίναξ παρατηρήσεων τῶν ἑρμηνευτικῶν μεταβλητῶν
β : εἶναι ἓν $N \times 1$ διαστάσεως διάνυσμα τῶν ἀγνώστων παραμέτρων
ε : εἶναι ἓν $T \times 1$ διαστάσεως διάνυσμα τῶν σφαλμάτων.

Ὑποθέσεις :

$E(\varepsilon) = 0$ ὅπου κατὰ Malinvaud [4] εἶναι ἀναγκαία ἡ ὑπόθεση αὕτη ἄλλως τὸ ὑπόδειγμα δὲν εἶναι ταυτοποιήσιμον.
$E(\varepsilon\varepsilon') = \sigma^2 I_\tau$
Ὁ πίναξ x ἔχει τάξιν N.

Ὁρισμὸς τοῦ προβλήματος τῆς πολυσυγγραμμικότητος

Ὅταν μεταξὺ τῶν ἀνεξαρτήτων μεταβλητῶν ὑφίστανται μία ἢ περισσότεραι γραμμικαὶ σχέσεις τότε λέγομεν ὅτι αἱ μεταβληταὶ αὐταὶ εἶναι πολυσυγγραμμικαί. (Δηλαδὴ ἰσχύει $\rho(x) = \alpha$ ὅπου $\rho(x)$ παριστᾶ τὴν τάξιν τοῦ πίνακος μὲ $0 \leqslant \alpha < N$).

Μία ἄλλη περίπτωσις εἰς τὴν ὁποίαν ἡ τάξις τοῦ πίνακος δὲν εἶναι ἴση μὲ N εἶναι ὅταν $T < N$· τὸ ὁποῖον σημαίνει ὅτι ὁ ἀριθμὸς τῶν παρατηρήσεων εἶναι μικρότερος ἀπὸ τὸν ἀριθμὸν τῶν συντελεστῶν ποὺ πρόκειται νὰ ἐκτιμηθοῦν. Τέτοια κατάστασις παρουσιάζεται ὅταν ἔχωμεν οἰκονομετρικὰ ὑποδείγματα, ἰδὲ F. M. Fisher [5]. Θὰ λέγωμεν ὅτι ἔχωμεν τελεῖαν πολυσυγγραμμικότητα ὅταν $|x'x| = 0$ ἐνῶ ὅταν συμβαίνει $|x'x| \rightarrow 0$ θὰ λέγωμεν ὅτι ἔχωμεν μὴ πλήρη πολυσυγγραμμικότητα τείνουσαν ὅμως πρὸς τὴν τελεῖαν πολυσυγγραμμικότητα. Εἰς τὴν πρᾶξιν παρουσιάζεται περισσότερον ἡ περίπτωσις τῆς μὴ πλήρους πολυσυγγραμμικότητος.

Δοθέντος, τώρα, τοῦ προβλήματος τῆς πολυσυγγραμμικότητος γεννᾶται ἡ ἀκόλουθος ἐρώτησις :

Τί πρέπει νὰ κάνωμεν ; Διὰ νὰ δώσωμεν ἀπάντησιν εἰς τὸ ἐρώτημα αὐτὸ θὰ ἀναφέρωμεν δύο λύσεις, ἡ μὲν μία ἡ ὁποία κάνει χρῆσιν τοῦ γενικευμένου ἀντιστρόφου πίνακος (Generalized Inverse Matrix Method), ἡ δὲ ἄλλη κάνει χρῆσιν τῆς Ridge Regression (RR).

2. Μέθοδος τοῦ γενικευμένου ἀντιστρόφου πίνακος (6,7,8,9)

Ἡ μέθοδος αὕτη κάνει χρῆσιν τοῦ γενικευμένου ἀντιστρόφου πίνακος διὰ τὴν ἀντιμετώπισιν τοῦ προβλήματος τῆς πολυσυγγραμμικότητος, οὕτω θεωροῦμεν σκόπιμον νὰ δώσωμεν ὡρισμένα στοιχεῖα ἐκ τῆς θεωρίας τῶν γενικευμένων πινάκων.

Ὁρισμός : Ὁ γενικευμένος ἀντίστροφος ἑνὸς $p \times q$ διαστάσεως πίνακος A οἱασδήποτε τάξεως εἶναι εἷς $q \times p$ πίναξ A^- ὁ ὁποῖος ἱκανοποιεῖ τὴν ἐξίσωσιν :

$$AA^- A = A \tag{2}$$

Εἰς τὴν βιβλιογραφίαν ὁ πίναξ A^- ἀναφέρεται καὶ ὡς: «-ἀντίστροφος», ἢ ὡς «ψευδοαντίστροφος» ἢ καὶ ὡς «ὑπὸ συνθήκην ἀντίστροφος».

Ἔχομεν νὰ παρατηρήσωμεν ὅτι ὁ πίναξ A^- δὲν εἶναι μονοσημάντως ὡρισμένος, ἤτοι ὑπάρχει ἄπειρον πλῆθος πινάκων A^- οἱ ὁποῖοι ἐπαληθεύουν τὴν (2). Εἶναι δὲ ὁ πίναξ A^- μονοσημάντως ὡρισμένος ὅταν $A^- = A^{-1} \leftrightarrow |A| \neq 0$.

Ὁ Penrose [6] ἀπέδειξεν ὅτι διὰ κάθε $p \times q$ πίνακα A ὑπάρχει εἷς καὶ μόνον εἷς $q \times p$ πίναξ A^+ τέτοιος ὥστε :

$$\begin{array}{ll} \text{(i)} & AA^+A = A \\ \text{(ii)} & A^+ AA^+ = A^+ \\ \text{(iii)} & (A^+A)' = A^+A \\ \text{(iv)} & (AA^+)' = AA^+ \\ \text{(v)} & trAA^+ = rankA = trA^+A \end{array}$$

Παρατηροῦμεν ἐκ τῆς (i) ὅτι, κάθε πίναξ τοῦ Penrose εἶναι ἐπίσης γενικευμένος ἀντίστροφος πίναξ.

Θὰ ἴδωμεν, τώρα, πῶς ἡ ὡς ἄνω θεωρία ἐφαρμόζεται εἰς τὴν περίπτωσιν ἐπιλύσεως συστημάτων $y = Ax$ ὅταν $|A| = 0$ μὲ x αἱ ζητούμεναι λύσεις. Κατὰ τὸν Rao [7, 8] διὰ τὸ σύστημα $y = Ax$ μὲ $|A| = 0$ ὅλαι αἱ λύσεις του δίδονται ἀπὸ τὴν σχέσιν $x^- = A^-y$.

Ἐπανερχόμεθα εἰς τὴν σχέσιν (1) τῆς πρώτης παραγράφου, ἀλλὰ μὲ τὸν πίνακα x μὴ ἱκανοποιοῦντα τὴν συνθήκην τάξεως, τότε ἐκ τῆς (1) ἔχομεν :

$$(x'x)\,\hat{\beta} = x'\,y \tag{3}$$

ἡ (3) δέ, δὲν ἔχει μονοσήμαντον λύσιν ἀφοῦ ὁ πίναξ x ἔχει τάξιν $\alpha < N$.

Διὰ νὰ λάβωμεν μίαν ἀπὸ τὰς λύσεις χρησιμοποιοῦμεν τὸν γενικευμένον ἀντίστροφον πίνακα $(x'x)^-$ τοῦ $(x'x)$ καὶ ἔχομεν τὴν λύσιν :

$$\tilde{\beta} = (x'x)^- x'y \tag{4}$$

Χρησιμοποιώντας ὅμως τὸν πίνακα Penrose ὁ ὁποῖος πλεονεκτεῖ λόγῳ τοῦ μονοσημάντου, μεταξὺ τῶν γενικευμένων πινάκων, ἔχομεν τὸν «ἄριστον — γραμμικὸν ἀμερόληπτον ἐκτιμητὴν τοῦ Penrose» τοῦ β ὁ ὁποῖος δίδεται ἀπὸ τὴν σχέσιν :

$$\beta^* = (x'x)^+ x'y \tag{5}$$

3. Ἡ μέθοδος τῶν Hoerl καὶ Kennard ([10, 11, 12, 13, 14]) ἢ ἡ μέθοδος τῆς Ridge Regression

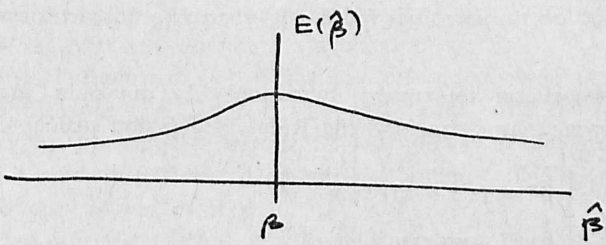
Πρὶν ἀπ' ὅλα, ἄς ἴδωμεν ποία εἶναι ἡ φιλοσοφία ἡ ὁποία διέπει τὴν μεθοδολογίαν τῶν Hoerl καὶ Kennard οὕτως ὥστε νὰ κατανοήσωμεν τὴν συμβολήν των εἰς τὸ πρόβλημα τῆς πολυσυγγραμμικότητος.

Καθὼς φαίνεται ἀπὸ τὴν σχέσιν Var $(\hat{\beta}) = \sigma^2 (x'x)^{-1}$ ἡ διακύμανσις γίνεται ἀρκούντως μεγάλη ὅταν ἔχωμεν πολυσυγγραμμικότητα μὲ ἀποτέλεσμα νά μὴ ἔχωμεν ἀξιόλογον ἐκτιμητὴν τοῦ β.

Συνεπῶς, κατὰ τοὺς Hoerl καὶ Kennard θὰ ἦτο δυνατὸν νὰ εὕρωμεν μίαν μεθοδολογίαν διὰ τῆς ὁποίας νὰ ἐλαχιστοποιεῖται ἡ διακύμανσις —αὐτὸ εἶναι τὸ κέρδος καὶ ἡ ἐπιδίωξις τῆς μεθοδολογίας τῶν Hoerl καὶ Kennard— καὶ ἀπὸ τὴν ἄλλην νὰ εἰσάγεται κάποια μεροληπτικότης.

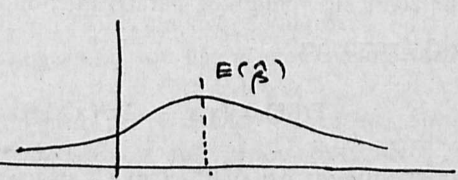
Γραφικῶς, δυνάμεθα νὰ παραστήσωμεν τὰς δύο περιπτώσεις, ἡ μία διὰ τὴν ὁποίαν ἐφαρμόζομεν τὴν τεχνικὴν τῶν ἐλαχίστων τετραγώνων καὶ μὲ τὴν παρουσίαν τοῦ προβλήματος τῆς πολυσυγγραμμικότητος καὶ ἡ ἄλλη ὅταν ἐφαρμόζομεν τὴν προταθεῖσαν ὑπὸ τοὺς Hoerl καὶ Kennard μεθοδολογίας.

Περίπτωσις πρώτη :



$E(\hat{\beta}) = \beta$ ἄρα καὶ μηδὲν μεροληπτικότητα, ἐνῶ ἡ τιμὴ τῆς Var$(\hat{\beta})$ εἶναι μεγάλη.

Περίπτωσις δευτέρα (ἡ ὑπὸ τῶν Hoerl καὶ Kennard) :



$E(\hat{\beta}) \neq \beta$ ἄρα ὑπάρχει κάποια μεροληπτικότης ἐνῶ ἡ τιμὴ τῆς Var$(\hat{\beta})$ εἶναι μικρή.

Μὲ μιὰ πρώτη ματιὰ δυνάμεθα νὰ εἴπωμεν ὅτι ἡ μεθοδολογία τῶν Hoerl καὶ Kennard παρουσιάζει τὸ πλεονέκτημα ὅτι χρησιμοποιεῖ ὅλες τὶς μεταβλητὲς αἱ ὁποῖαι περιλαμβάνονται εἰς τὸ οἰκονομετρικὸν ὑπόδειγμα (π.χ. χωρὶς τὴν ἀφαίρεσιν ὡρισμένων μεταβλητῶν) μὲ μόνον μειονέκτημα ὅτι ἔχομεν μεροληπτικὸν ἐκτιμητήν.

Οὕτως οἱ Hoerl καὶ Kennard ἀντὶ τοῦ $\hat{\beta} = (x'x)^{-1}x'y$ προτείνουν τὸν ἀκόλουθον ἐκτιμητήν.

$$\hat{\beta}^* = (x'x + \kappa I)^{-1}x'y \tag{1}$$

μὲ $\kappa \geqslant 0$ εἰς τὸν ὁποῖον ἔδωσαν τὴν ὀνομασίαν Ridge Regression.

Ἐκ τῆς (1) εἶναι προφανὲς ὅτι διὰ $\kappa = 0$ ὁ ὡς ἄνω ἐκτιμητὴς μεταπίπτει εἰς τὸν γνωστὸν ἐκτιμητὴν τῶν ἐλαχίστων τετραγώνων.

Ἡ (1) εἰς τὴν γενικευμένην της μορφὴν γράφεται :

$$\hat{\beta}^* = (x'x + KI)^{-1}\, x'y \tag{2}$$

ὅπου ὁ K εἶναι εἷς διαγώνιος πίναξ.

Ἀμέσως, γίνεται ἀντιληπτὸν ὅτι τὸ πρόβλημα ἐδῶ εἶναι τὸ πῶς θὰ εὕρωμεν τὸ κ (μειονέκτημα τῆς μεθόδου).

Οἱ Hoerl καὶ Kennard προτείνουν διαφόρους τρόπους εὑρέσεως τοῦ κ· ὁ ἁπλούστερος δὲ τούτων εἶναι ἡ διὰ τῆς γραφικῆς παραστάσεως μέθοδος τοῦ $\hat{\beta}^*$ ἔναντι τοῦ κ.

3.1. Ἀξιοσημείωτοι ἰδιότητες: ἀναφέρομεν ἐν συντομίᾳ τὰς πλέον βασικὰς ἰδιότητας τῶν ἐκτιμητῶν τῆς Ridge Regression μεθόδου.

α) Ἐκ τῶν $\left\{\begin{array}{l}\hat{\beta}^* = (x'x+ \kappa I)^{-1}x'y \\ \hat{\beta} \;\, = (x'x)^{-1}x'y\end{array}\right.$

ἕπεται ὅτι :

$$\hat{\beta}^* = (x'x + \kappa I)^{-1}x'y = (x'x + \kappa I)^{-1}\,(x'x)\,\hat{\beta} \tag{α1}$$

ἤτοι ὁ $\hat{\beta}^*$ εἶναι εἷς γραμμικὸς μετασχηματισμὸς τοῦ β.

β) Ἐκ τῆς (α1) ἕπεται ὅτι

$$E(\hat{\beta}^*) = (x'x + \kappa I)^{-1}(x'x)\,\beta \neq \hat{\beta}^*$$

ἐκ τῆς ὁποίας ἐμφαίνεται ὅτι ὁ $\hat{\beta}^*$ εἶναι εἷς μεροληπτικὸς ἐκτιμητής.

γ) Ἡ διακύμανσις δίδεται ἐκ τῆς σχέσεως :

$$\mathrm{Var}\,(\hat{\beta}^*) = \sigma^2\,(x'x + \kappa I)^{-1}\,(x'x)\,(x'x + \kappa I)^{-1}.$$

δ) Ἀποδεικνύεται ὅτι ἡ διακύμανσις εἶναι φθίνουσα συνάρτησις τοῦ κ.

ε) Ἀποδεικνύεται ὅτι ἡ μεροληπτικότης εἶναι αὔξουσα συνάρτησις τοῦ κ.

3.2. Τί ἔδειξαν αἱ μέχρι τοῦδε ἐφαρμογαὶ

Κατὰ τοὺς Brown καὶ Beattie ([15]) ἡ μέθοδος τῶν Hoerl καὶ Kennard ἀπαιτεῖ a priori πληροφορίας διὰ τὴν ἐπιτυχῆ ἐφαρμογήν της. Εἰδικώτερον διὰ τὰ οἰκονομικὰ ὑποδείγματα προκύπτουν τὰ καλύτερα ἀποτελέσματα ἐφ' ὅσον ὅλοι οἱ συντελεσταὶ τῆς παλινδρομήσεως ἔχουν τὸ αὐτὸ πρόσημον καὶ στατιστικῶς εἶναι ἀξιόλογοι.

Κατὰ τοὺς Dempster, Schatzoff καὶ Wermuth ([16]) εἶναι δυνατὸν νὰ ἔχωμεν βελτιώσεις μέσῳ τῆς τεχνικῆς τῶν Hoerl καὶ Kennard ὅταν ἡ σχέσις ἐξαρτήσεως μεταξὺ τῶν ἑρμηνευτικῶν μεταβλητῶν εἶναι ἰσχυρά.

Ὁ Farebrother ([17]) ἔδειξε τὴν ὁμοιότητα τῆς Ridge Regression τεχνικῶς μὲ ἐκείνης τῆς ἐλαχιστοποιήσεως τοῦ μέσου σφάλματος τετραγώνου.

Ὁ Hsiang ([18]) ἔδειξεν ὅτι κάτω ἀπὸ ὡρισμένες ὑποθέσεις οἱ ἐκτιμηταὶ διὰ τῆς μεθόδου τῶν Hoerl καὶ Kennard εἶναι ἰσοδύναμοι μὲ τοὺς Bayesian ἐκτιμητάς.

Οἱ Lindley καὶ Smith ([19]) χρησιμοποιώντας τὴν θεωρίαν τοῦ Bayes εἰς τὸ γραμμικὸν ὑπόδειγμα ἔλαβον ἐκτιμήσεις ὁμοίας μὲ ἐκείνας τῶν Hoerl καὶ Kennard μὲ τὴν μόνην διαφορὰν ὡς πρὸς τὴν ἐκλογὴν καὶ τὴν ἑρμηνείαν τοῦ κ.

Ὁ Mcdonald ([20]) ἔδειξεν ὅτι οἱ συντελεσταὶ τῆς Ridge Regression δύνανται νὰ ἐκφρασθοῦν ὡς ρηταὶ συναρτήσεις τῆς παραμέτρου κ.

Οἱ McDonalt καὶ Galarman ([21]) ἀπέδειξαν διὰ παραδειγμάτων ὅτι ὅταν αἱ μεταβληταὶ προβλέψεως εἶναι ἰσχυρῶς συσχετισμέναι, τότε ἡ τεχνικὴ τῆς Ridge Regression εἶναι προτιμητέα ἐκείνης τῶν ἐλαχίστων τετραγώνων.

Οἱ Smith καὶ Goldstein ([22]) κριτικάρωντας τὸ ἄρθρον τῶν Conniffe καὶ Stone ([23]) εἰς τὸ ὁποῖον ἰσχυρίζονται ὅτι :

(α) οἱ ἐκτιμηταὶ διὰ τῆς μεθόδου τῆς Ridge Regression δὲν ἔχει ἀποδειχθῆ ὅτι ἔχουν μικρότερον μέσον σφάλμα τετραγώνου διὰ ἐκτιμητὰς μὲ τὸ κ ἐκτιμηθέν.

(β) Ἀφοῦ τὸ κ ἔχει ἐκτιμηθῆ, αἱ διακυμάνσεις τῶν ἐκτιμητῶν τῶν συντελεστῶν δὲν μποροῦν νὰ ὑπολογισθοῦν.

(γ) Τὰ κριτήρια διὰ τὴν ἐκλογὴν τοῦ κ δὲν εἶναι ἀκριβῆ.

Ἔναντι τούτων οἱ Smith καὶ Goldstein ἀναφέρουν ὅτι :

(α΄) ὑπάρχει δυνατότης μειώσεως τοῦ μέσου σφάλματος τετραγώνου διὰ μικρὲς τιμὲς τοῦ κ,

(β΄) τὸ κ δὲν εἶναι συνάρτησις τοῦ y καὶ ὡς ἐκ τούτου δύναται νὰ θεωρηθῆ ὡς σταθερὰ κατὰ τὸν ὑπολογισμὸν τοῦ μέσου σφάλματος τετραγώνου,

(γ΄) ἡ τεχνικὴ τῆς Ridge Regression βελτιώνει τὸ μέσον τετραγωγικὸν σφάλμα,

(δ΄) ἡ Ridge Regression εἶναι προτιμητέα ἔναντι τῶν ἐλαχίστων τετραγώνων ὅταν δὲν ὑπάρχουν ἀρκετὰ στατιστικὰ στοιχεῖα.

4. Παρατηρήσεις ἐπὶ τοῦ ἄρθρου τοῦ Σακέλλη

(i) Ἔχομεν νὰ παρατηρήσωμεν διὰ τήν «μέθοδον τῆς συγκρίσεως τῶν μερικῶν συντελεστῶν, μὲ τὸν ὁλικὸ συντελεστὴ πολλαπλοῦ προσδιορισμοῦ» ὅτι :

(α) ἡ ὑπὸ τοῦ Klein μέθοδος εἶναι «Rule of Thumb»

(β) δὲν εἶναι καὶ ἀναγκαία συνθήκη

(γ) ἰδὲ παρατήρησιν (iii) κατωτέρω.

(ii) Εἰς τὴν σελίδα 889 διαβάζωμεν ὅτι «μὲ τὴν εἰσαγωγὴ μιᾶς μεταβλητῆς ....» χωρὶς νὰ ὁρίζεται ποία εἶναι ἡ μεταβλητὴ καθὼς καὶ ἀνάλογες στατιστικὲς ὑποθέσεις εἰς τὶς ὁποῖες πρέπει νὰ ὑπακούη ἡ μεταβλητή.

Ἐπὶ πλέον διὰ νὰ μελετήσωμεν τὴν μορφὴν τῆς πολυσυγγραμμικότητος βάσει τῆς στατιστικῆς t προϋποτίθεται ὅτι τὰ x_i καὶ x_j (σελὶς 889 πρᾶγμα ποὺ δὲν ἀναφέρεται) εἶναι μεταξύ τους ἀνεξάρτητες καὶ ἀκολουθοῦν τὴν κανονικὴν κατανομήν· αὐτὸ δὲ τὸ κριτήριον ἐφαρμόζεται διὰ κάθε ζεῦγος x_i καὶ x_j.

(iii) Ἀφοῦ ὁ Σακέλλης παρουσιάζει μίαν περίληψιν τῆς φιλολογίας τῆς πολυσυγγραμμικότητος θὰ ὤφειλε πρὸς πλήρη ἐνημέρωσιν τῶν ἀναγνωστῶν νὰ παρουσιάση τὰ μέχρι σήμερον γραφέντα ἐπὶ τοῦ θέματος ἀπόψεις καὶ κρίσεις. Ἐπὶ παραδείγματι, θὰ ἦτο χρησιμώτατον νὰ πληροφορηθῇ ὁ ἀναγνώστης τὶς ἀπόψεις τοῦ R. Wichers ([24]) διὰ τὸ τρίτον στάδιον τῆς μεθοδολογίας τῶν Farrar καὶ Glauber — ἤτοι τὴν εὕρεσιν τῆς μορφῆς τῆς πολυσυγγραμμικότητος.

Εἰς τὸ ἄρθρον του ὁ P. Wichers ἀπέδειξεν ὅτι τὸ τρίτον στάδιον τοῦ κριτηρίου τῶν Farrar καὶ Glauber εἶναι μὴ ἀποτελεσματικόν. Κατ' ἀρχὰς θεωροῦμεν τὸν συσχετισμένον πίνακα Π τοῦ γραμμικοῦ ὑποδείγματος καὶ μὲ π_i παριστῶμεν τὰς στήλας του.

Ποία ἦτο ἡ στρατηγικὴ τοῦ Wichers ; Διὰ νὰ ἀποδείξη ὅτι τὸ τρίτο μέρος τοῦ κριτηρίου τῶν Farrar καὶ Glauber δὲν ἰσχύει, κατεσκεύασε δύο συσχετισμένους πίνακες οἱ ὁποῖοι ἔδιδον τὴν αὐτὴν τιμὴν εἰς τὸν μερικὸν συντελεστὴν r_{ij} καὶ μὲ διαφορετικὲς μορφὲς πολυσυγγραμμικότητος. Μάλιστα δὲ εἰς τὸ παράδειγμά του προέκυψαν τὰ ἀκόλουθα ἀποτελέσματα :

α) Διὰ τὸν πρῶτον πίνακα : $r_{12} = -2^{-1/2}$ καὶ $\pi_1' \pi_2 = .99$ καὶ ὁπότε αἱ στῆλαι π_1 καὶ π_2 παρουσιάζουν ἰσχυρὰν συσχέτισιν.

β) Διὰ τὸν δεύτερον πίνακα : $r_{12} = -2^{1/2}$ καὶ $\pi_1'\pi_2 = 0$ καὶ συνεπῶς αἱ στῆλαι π_1 καὶ π_2 εἶναι ὀρθογώνιαι.

Ἄρα τὸ συμπέρασμα τοῦ Wichers εἶναι ὅτι τὸ r_{ij} δὲν εἶναι τὸ κατάλληλον μέτρον τῆς πολυσυγγραμμικότητος, ταυτόχρονα προτείνει μίαν νέαν στατιστικὴν ἡ ὁποία κατανέμεται ὡς ἡ F.

(iv) Τὰ κριτήρια τὰ ὁποῖα ἀνέπτυξαν οἱ Farrar καὶ Glauber στηρίζονται ἐπὶ τῆς ὑποθέσεως ὅτι αἱ ὑπὸ μελέτην μεταβληταὶ ἀκολουθοῦν μίαν κανονικὴν κατανομήν· αὐτὴ εἶναι μία πολὺ λεπτὴ ὑπόθεσις καὶ τὸ διατὶ μᾶς τὸ δίδουν τὰ ἄρθρα τῶν Kumar ([25]) καὶ O'Hagan - McCabe ([26]).

Ὁ Kumar ὀρθῶς διακρίνει δύο μορφὲς γραμμικῶν ὑποδειγμάτων ἐξαρτωμένων ἀπὸ τὰς ὑποθέσεις αἱ ὁποῖαι γίνονται ὅσον ἀφορᾶ τὸν πῖνακα x, ἤτοι ἂν εἶναι στοχαστικὸς ἢ ὄχι.

Ἐὰν ὁ πίναξ x δὲν εἶναι στοχαστικὸς τότε κατὰ τὸν Kumar «τὸ πρόβλημα τῆς πολυσυγγραμμικότητος δύναται νὰ θεωρηθῇ ὡς ἀριθμητικὸν πρόβλημα τὸ ὁποῖον δὲν ὑπόκειται εἰς στατιστικὰ κριτήρια».

Ἀφοῦ πλέον τὸ θέμα τῆς πολυσυγγραμμικότητος, ὅταν ὁ πίναξ x εἶναι μὴ στοχαστικός, ἀπὸ στατιστικὸν πρόβλημα μεταπίπτει εἰς ἀριθμητικὸν γεννᾶται αὐτομάτως ἡ περίπτωσις τῆς διερευνήσεως ὅταν ὁ πίναξ x εἶναι στοχαστικός.

Ὁ Kumar ἐξετάζει τὴν ἰσχὺν τῶν στατιστικῶν ἀποτελεσμάτων τοῦ κριτηρίου τῶν Farrar καὶ Glauber εἰς τὴν περίπτωσιν καθ' ἣν ὁ πίναξ εἶναί στοχαστικός

Καὶ νὰ ποῦ καταλήγει τὸ ἀξιόλογον ἄρθρον τοῦ Kumar :

«Ὁ κύριος σκοπὸς τοῦ παρόντος σημειώματος εἶναι νὰ ἐπισυνάψω ἀμφιβολία ἐπὶ τοῦ x^2 κριτηρίου διὰ τὴν ὕπαρξιν τῆς πολυσυγγραμμικότητος καὶ ἐπὶ τῶν F καὶ t κριτηρίων διὰ τὸν προσδιορισμὸν τῆς θέσεως τῆς πολυσυγγραμμικότητος ... Ἀλλὰ ἡ ἀπόφασις τὸ πῶς αὐτὰ τὰ μέτρα πρέπει νὰ χρησιμοποιοῦνται δὲν βασίζεται ἐπὶ γνωστῆς θεωρίας. Προσέτι, τὰ κριτήρια x^2, F καὶ t τὰ ὁποῖα προτείνονται ὑπὸ τῶν Farrar καὶ Glauber δὲν δύνανται οὔτε νὰ προσεγγισθοῦν ἀφοῦ ἡ βασικὴ θεωρία κατανομῆς εἶναι ἐσφαλμένη.

Ἅπαξ τὰ ἀποτελέσματα τῶν κατανομῶν τῆς ἐργασίας των ἔχουν ἀπορριφθῆ, τὰ τρία μέτρα τὰ ὁποῖα προτείνουν δὲν προσφέρουν περαιτέρω ἐμβάθυνσιν τοῦ προβλήματος. Καὶ ὀρθῶς οἱ O'Hagan καὶ McCabe καταλήγουν «Οὕτω, εἴμεθα βέβαιοι ὅτι δὲν δυνάμεθα νὰ .... συμφωνήσωμεν ἐπὶ τοῦ σημειώματος τῆς νίκης, ὅτι τὸ πρόβλημα ἔχει λυθῆ καὶ δὲν εἶναι ἀναγκαία περαιτέρω ἐπανεξέτασις (Farrar καὶ Glauber, σελ. 10)».

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Σακέλλη Μ. : «Τὸ πρόβλημα τῆς πολυσυγγραμικότητος στὴν Οἰκονομετρία», Σπουδαί, τόμος ΚΣΤ', τεῦχος 4, 1976.
2. Giles, D.E.A. : Essays on Econometric Topics : from theory to practice, Reserve Bank of New Zealand, 1973.
3. D.E. Farrar and R.R. Glauber : «Multicollinearity in Regression Analysis : The Problem Revisited», R.E. Stat. 1966.
4. Malinvaud, E. : Statistical Methods of Econometrics, North-Holland 1970.,
5. Fisher, F.M. «Dynamic Structure and Estimation in, Economy-wide Econometric Models», in J.S. Duesenberry et al. The Brookings Quarterly Econometric Model of the United States. Rand-McNally, 1965.
6. Penrose, R.A. : «A Generalized Inverse for Matrices», Proc. Cambridge Philos. Soc., 51, 1955.
7. Rao, C.R. : «A Note on Generalized Inverse of a Matrix with Applications to Problems in Mathematical Statistics», J.R.S. Soc. Series B, 24.
8. Rao, C.R. : Linear Statistical Inference and its Application, Wiley, 1965.
9. Neeleman, D. : Multicollinearity in Linear Models Tilburg University Press, 1973.

10. Hoerl, A.E. : Application of Ridge Analysis to Regression Problems». Chemical Engineering Progress, 1962.

11. Hoerl, A.E. and Kennard, P.W. : «On Regression Analysis and Biased Estimation», 127th annual meeting of the American Statistical Association.

12. Hoerl, A.E. and Kennard, R.W. : Ridge Regression : Application to Non-Orthogonal Problems. Technometrics, 12, 1970.

13. Hoerl, A.E. and Kennard, R.W. : «Ridge Pegression : Biased Estimation for Nonorthogond Problems», Technometrics, 12, 1970.

14. Hoerl, A.E. and Kennard, R.W. : «A Note on Power Generalization of Ridge Regression», Technometrics, 17, 1975.

15. Brown, W.G. and Beattie, B.R. : «Improving Estimates of Economic Parameters by Use of Ridge Regression with Production Function Applications» A.J. of Agricultural Economics, 57, 1975.

16. Dempster, A.P., Schatzoff M. and Wermuth, N. : «It Simulation Study of Alternatives to OLS» JASA, 1976.

17. Farebrother, R.W. : «The Minimum Mean Square Error Linear Estimator and Ridge Regression», Technometrics, 17, 1975.

18. Hsiang, T.C. : «Bayesian View of Ridge, Regression» Statistician, 24, 1975.

19. Lindley, D.V. and Smith, A.F.M. : «Bayes Estimates for the Linear Model» J.R.S. Soc. Series B, 34, 1972.

20. McDonald, G.C. : «Some Properties of Ridge Coefficients», *General Motors Research Publications*, 1975.

21. McDonald, G.C. and Galarman, D.I. : «Monte Carlo evaluation of Some Ridge - Type Estimators» G.M.R.P. 1973.

22. Smith A.F.M. and Goldstein, M. : «Ridge Regression : Some Comments on a Paper of Conniffe and Stone» Statistician, 24, 1975.

23. Conniffe, D. and Stone J. : «A Critical View of Ridge Regression», Statistician, 22, 1973.

24. Wichers, R.C. : «The Detection on Multicollinearity : A comment», R.E. Stat. 1975.

25. Kumar K.T. : «Multicollinearity in Regression Analysis» R.E. Stat. 1975.

26. O'Hagan, J. and McCabe. : «Tests for the Severity of Multicollinearity in Regression Analysis : A Comment», R.E. Stat. 1975.