AD HOC ΕΠΙΛΥΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ ΤΗΣ ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑΣ: Η RIDGE ΠΑΛΙΝΔΡΟΜΗΣΗ

Υπό
ΚΩΣΤΑ Π. ΣΥΡΙΟΠΟΥΛΟΥ
Πανεπιστήμιο Aix - Marseille III - ΓΑΛΛΙΑ

ΠΕΡΙΛΗΨΗ

Το πρόβλημα της πολυσυγγραμμικότητας είναι γνωστό στην οικονομετρία. Και τούτο γιατί είναι δύσκολο να βρούμε στα οικονομετρικά υποδείγματα μεταβλητές τελείως ανεξάρτητες μεταξύ τους. Επειδή όμως η εκτίμηση ενός υποδείγματος, όπου παρουσιάζεται αυτό το πρόβλημα, με την Μέθοδο Ελαχίστων Τετραγώνων (Μ.Ε.Τ.) δίνει αποτελέσματα ασταθή και αναξιόπιστα, διάφορες μέθοδοι επίλυσης έχουν, κατά καιρούς, προταθεί. Στο παρόν άρθρο θα εξετάσουμε την Ridge Regression (R. R.), η οποία φαίνεται να είναι πιο κοντά στην πραγματικότητα. Παρουσιάζουμε την R.R. στην παράγραφο II μετά από μια σύντομη υπενθύμιση της Μ.Ε.Τ. και του προβλήματος της πολυασυγγραμμικότητας. Στην παράγραφο III εκθέτουμε τις διάφορες ερμηνείες της μεθόδου αυτής και στην παράγραφο IV μία προέκταση της R.R. στα συστήματα αλληλοεξαρτώμενων εξισώσεων. Στην παράγραφο V δοκιμάσαμε μία εφαρμογή της Ridge σε ένα υπόδειγμα, όπου παρουσιάζεται το πρόβλημα της πολυσυγγραμμικότητας.

*. Θέλω να ευχαριστήσω τον Καθηγητή J,P. Marciano (Aix - Marceille III) και τον Καθηγητή Α. Ηλιάδη (Aix - Marseille II), για την βοήθεια τους στην εφαρμογή αυτή. Επίσης τον Καθηγητή Α. Λαζαρίδη (Α.Π.Θ.) για τις χρήσιμες παρατηρήσεις του σε προγενέστερη μορφή του άρθρου αυτού.

I. Η Μ.Ε.Τ.— ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ

Έστω, το προς εκτίμηση υπόδειγμα :

(1) $$Y = X\beta + U$$

όπου,

Y = είναι ένα διάνυσμα στήλη (nx1), X μία μήτρα (nxp), β το διάνυσμα των συντελεστών της παλινδρόμησης (px1), U το διάνυσμα στήλη (nx1) του διαταρακτικού όρου, με μέσο μηδέν και $E(UU') = \sigma^2 I_n$, n το μέγεθος του δείγματος και p το πλήθος των παραμέτρων.

Με την Μ.Ε.Τ., ως γνωστόν, ελαχιστοποιούμε το άθροισμα των τετραγώνων των καταλοίπων S :

2) $$s = \sum_i \hat{U}_i^2 = ||Y - X\hat{\beta}||^2 = Y'Y - 2\hat{\beta}'(X'Y) + \hat{\beta}'(X'X)\hat{\beta}$$

Από το πρόγραμμα της ελαχιστοποίησης και αν η αντίστροφη μήτρα $(X'X^{-1}$ υπάρχει, παίρνουμε ι

(3) $$\hat{\beta} = (X'X)^{-1} X'\gamma$$

Αποδεικνύεται δε ότι ο εκτιμητής $\hat{\beta}$ είναι αμερόληπτος και με την μικρότερη διακύμανση (B.L.U.E.).

Ο ρ ι σ μ ό ς τ η ς π ο λ υ σ υ γ γ ρ α μ μ ι κ ό τ η τ α ς (2,8). Για να είναι η σχέση (3) μοναδική λύση της (2), πρέπει η (2) να είναι αυστηρά κυρτή. Αυτό σημαίνει ότι η Εσσιανή μήτρα της (2) πρέπει να είναι θετικά ορισμένη :

(4) $$H = (\theta^2 S/\theta\hat{\beta}\theta\hat{\beta} = 2X'X$$

Για να είναι η μήτρα $X'X$ θετικά ορισμένη, πρέπει η μήτρα X να έχει πλήρη βαθμό. Αν ο βαθμός της X είναι ίσος με p, τότε δεν υπάρχει γραμμική εξάρτηση. Αντίθετα, όταν $r(X)<p$, τότε υπάρχει γραμμική εξάρτηση και η μήτρα $X'X$ δεν είναι αντιστρέψιμη.

Για να εξετάσουμε την πολυσυγγραμμικότητα, ορίζουμε την ευκλείδια τετραγωνική απόσταση μεταξύ $\hat{\beta}$ και β :

$$(5), \qquad L^2 = (\hat{\beta} - \beta)' (\hat{\beta} - \beta)$$

με μαθηματική ελπίδα :

$$(6) \qquad E(L^2) = \sigma^2 \operatorname{tr} (X'X)^{-1} = \sigma^2 \sum_i (1/\lambda_i),$$

όπου, λ_i είναι οι χαρακτηριστικές ρίζες της μήτρας $(X'X)$ και διακύμανση

$$(7) \qquad Var(L^2) = 2\sigma^4 \sum_i (1/\lambda_i) \quad \text{με } \lambda_{max} = \lambda_1 \geqslant \ldots \geqslant \lambda_p = \lambda_{min} > 0$$

Στην περίπτωση που παρουσιάζεται πολυσυγγραμμικότητα η μήτρα $(X'X)$ αποκλίνει της ταυτοτικής I_p (που είναι η ορθογώνια μορφή της) και περιέχει μια τουλάχιστον χαρακτηριστική ρίζα που τείνει στο μηδέν. Έτσι η μέση τετραγωνική απόσταση μεταξύ $\hat{\beta}$ και β (σχέση 6) θα είναι σχετικά μεγάλη.

Για την αποκάλυψη και τον έλεγχο της πολυσυγγραμμικότητας παραθέτουμε τον κάτωθι πίνακα :

Έχουμε πολυσυγγραμμικότητα όταν:	Έλεγχοι
1) $\lvert X'X \rvert \longrightarrow 0$	I) Το Test των FARRAR-GLAMEER
2) $r_{ij} \longrightarrow 1$ (1)	$H_0: \lvert X'X \rvert = 1$ έναντι της $H_a: \lvert X'X \rvert < 1$ $\chi^2(v) \simeq x \cdot \log \lvert X'X \rvert,\ x = -[n-1-\frac{1}{6}(2p+5)]$
3) μία τουλάχιστον χαρακτηριστική ρίζα της $(X'X)$ τείνει στο μηδέν	II) Το Test του HAITOVSKY $H_0: \lvert X'X \rvert = 0$ έναντι της $H_a: 0 < \lvert X'X \rvert < 1$
4) $\sigma^2\ V.I.F(j) \longrightarrow \infty$	$\chi^2(v) \simeq x \cdot \log(1-\lvert X'X \rvert)$
5) $R_c = \frac{V.I.F.(j)}{P} > 1$	III) Tests κατά BAYES . του ορθογωνίου . της γωνίας . της ζημίας

Ονομάζουμε V.I.F. (Variance Inflated Factor)), σχετικά με την εξωγενή μεταβλητή j, την έκφραση :

$$V.I.F.\,(j) = q_{jj}^{-1} = (1-\dot{R}_j^2\, j)^{-1},$$

όπου, $q^{-1} = P.\Lambda^{-1}, P'$, όπου P η μήτρα των χαρακτηριστικών διανυσμάτων και Λ=διαγ. $(\lambda_1, \ldots, \lambda_p)$

Όταν $R_j^2 \longrightarrow 1$ σημαίνει την ύπαρξη της πολυσυγγραμμικότητας και το q_{jj}, θα είναι μεγάλο, όπως επίσης και η διακύμανση του εκτιμητή $\hat{\beta}_j$, γιατί Var $(\hat{\beta}_j) = q_{jj}\sigma^2$. Με άλλα λόγια ισχύει η σχέση :

$$\text{Var}\,(\hat{\beta}_j) = q_{jj}\sigma^2 > \sigma^2,\ \text{αφού}\ q_{jj} > 1,\ \text{όταν}\ R_j^2 > 0.$$

Κατά συνέπεια το q_{jj} μετράει το «φούσκωμα» των εκτιμητών της Μ.Ε.Τ. οφειλόμενο στην πολυσυγγραμμικότητα.

---

1. Ας σημειώσουμε ότι η μεγάλη συσχέτιση μπορεί να χρησιμοποιηθεί σαν δείκτης ύπαρξης της πολυσυγγραμμικότητας, αλλά η απουσία μεγάλης συσχέτισης δε σημαίνει και απουσία πολυσυγγραμμικότητας (βλ. P. Dhrymes : Introductory Econometrics, 1978, σελ. 191). Ένας άλλος πρακτικός τρόπος είναι να δούμε $R_{12}^2 > R_{\cdot Y.\,12}^2$, σύμφωνα με τον Klein (βλ. (2) σελ. 12 - 13.

Είδαμε ότι : $E(L^2) = \sigma^2 tr (X'X)^{-1}$ και σύμφωνα με τα ανωτέρω θα έχουμε :

$$E(L^2) = \sigma^2 n^{-1} \sum V.I.F. (j)$$

Αν $R_j^2 = 0$ τότε VIF (j) = 1 και η ανωτέρω σχέση γίνεται : $E(L^2) = n^{-1}.\sigma^2.p$.

Μπορούμε λοιπόν να θεωρήσουμε σαν δείκτη της ύπαρξης της πολυσυγγραμμικότητας τη σχέση (βλ. (5) :

$$Rc = \frac{Tr \quad (\Sigma_{\hat{\beta}\hat{\beta}})}{\frac{1}{n} \cdot \sigma^2 . p} = \frac{\frac{1}{n} \cdot \sigma^2 \cdot \sum_{j=1}^{p} (VIF)j}{\frac{1}{n} \cdot \sigma^2 \cdot p} = \frac{\sum_j (VIF)j}{p}$$

όπου

$\sum \hat{\beta} \hat{\beta}$ η μήτρα των διακυμάνσεων - συνδιακυμάνσεων.

$R_c = 1$, τότε οι εξωγενείς μεταβλητές είναι ανεξάρτητες

$R_c > 1$, τότε υπάρχει πρόβλημα πολυσυγγραμικότητας

Ας σημειωθεί ακόμα (βλ. (2, 13) ότι οι συγγραφείς Kumar (1975), O Hagan και McCabe (1975) και ο G. S. Maddala (1977) κριτίκαραν τους ελέγχους των Farrar - Glauber και του Haitovsky, υποστηρίζοντας ότι η πολυσυγγραμμικότητα αναφέρεται στο δείγμα και δεν έχει νόημα να ελέγξουμε την υπόθεση άν το φαινόμενο παρουσιάζεται ή όχι στον πληθυσμό. Τέλος, ο κλασσικές μέθοδοι επίλυσης του προβλήματος (αφαίρεση εξωγενών μεταβλητών που παρουσιάζουν ισχυρή συσχέτιση, άθροιση αυτών, χρησιμοποίηση εξωγενών πληροφοριών κλπ.) δεν φαίνονται ικανοποιητικές. Άλλες μέθοδοι είναι : η μέθοδος του γενικευμένου αντίστροφου πίνακα (βλ. και (8), η ανάλυση των κυριοτέρων συνιστωσών και η παλινδρόμηση Riddge με την οποία θα ασχοληθούμε στη συνέχεια.

II. Η ΠΑΛΙΝΔΡΟΜΗΣΗ RIDGE : ΚΛΑΣΙΚΗ ΚΑΙ ΓΕΝΙΚΕΥΜΕΝΗ

Η κεντρική ιδέα της μεθόδου αυτής είναι να εισαγάγουμε στη μήτρα συσχετίσεως (X'X) των εξωγενών μεταβλητών ένα διαταρακτικό όρο ή σταθερά μεροληψίας K και να εξετάσουμε την ευαισθησία των εκτιμητών της παλινδρόμησης στις μεταβολές του συντελεστή K, παρουσία της πολυσυγγραμμικότητας στο υπόδειγμα μας. Στην περίπτωση αυτή οι εκτιμητές Ridge είναι μεροληπτικοί αλλά για μερικές τιμές του K μπορούμε να πάρουμε εκτιμητές πιο σταθερούς και πιο αποτελεσματικούς από αυτούς της Μ.Ε.Τ., στο επίπεδο του Μέσου Τετραγωνικού Σφάλματος (Μ.Τ.Σ.).

Το μειονέκτημα της μεθόδου αυτής είναι ότι η εκλογή της άριστης τιμής του K εξαρτάται από την παράμετρο β και τη δια.κύμαση σ^2. Λόγω της δυσκολίας αυτής διάφοροι συγγραφείς πραγματοποίησαν πολλές μεθόδους εξομοίωσης (Simulations) Monte Carlo ή άλλες με σκοπό να εξετάσουν τις ιδιότητες των εκτιμητών Ridge. Και 7ΐάλι όμως δεν μπορούμε να βγάλουμε οριστικά συμπεράσματα γιατί κάθε αποτέλεσμα εξαρτάται από το συγκεκριμένο υπόδειγμα και τη δομή του, το μέγεθος του δείγματος, το βαθμό της πολυσυγγραμμικότητας κλπ. ([1])

1) Η κλασική μέθοδος Ridge (O.R.R.) (6) :

Η οικογένεια των εκτιμητών Ridge δίνεται από τη σχέση :

$$\hat{\beta}\,(K) = (X'X + K,I_p)^{-1}\, X'Y, \tag{8}$$

όπου K είναι μη στοχαστικό και $K \in (0,1)$.

Κυριώτερες ιδιότητες:

α) $$\hat{\beta}\,(K) = (X'X + KI_p)^{-1}\,(X'X)\,(X'X)^{-1}\,X'Y = (X'X + KI_p)^{-1}\,X'X\hat{\beta} =$$
$$= [I_p + K\,(X'X)^{-1}]^{-1}\,\hat{\beta} = Z_k\beta \tag{9}$$

δηλαδή $\hat{\beta}\,(K)$ είναι γραμμικός μετασχηματισμός του $\hat{\beta}\,(0)$ και όπου

$$Z_k = (X'X + KI_p)^{-1}\,(X'X) = [X'X\{\,I_p + K\,(X'X)^{-1}\,\}]^{-1}\,(X'X) =$$

1. Υπάρχουν όμως περιπτώσεις όπου η εγγύηση της μείωσης του Μ. Τ. Σ. για παράδειγμα γίνεται και χωρίς την ανάλυση τέτοιων μελετών (βλ. για παράδειγμα (3).

$$= (I_p + K(X'X)^{-1})^{-1}(X'X)^{-1}(X'X) = (I_p + K(X'X)^{-1})^{-1} := Z', \text{ και}$$

$$Z' = (X'X + K I_p)^{-1}(X'X) . \; Z'\ (X'X) + KI_p) = X'X \text{ ή}$$

$$Z'\ (X'X) + KZ' = (X'X) \text{ ή } Z'(X'X) = X'X - KZ'$$

$$Z' = I_p - KZ'(X'X)^{-1} = (I_p - K(X'X + KI_p))(X'X)(X'X)^{-1} = I_p - K(X'X + KI_p)^{-1}$$

Κατά συνέπεια έχουμε:

$$Z = Z' = (X'X + KI_p)^{-1}(X'X) = [I_p - K(X'X)^{-1}]^{-1} = I_p - K(X'X + KI_p)^{-1}$$

Από τη σχέση (8) είναι φανερό ότι για $K = 0$ βρίσκουμε τον εκτιμητή της Μ.Ε.Τ. Μ.Ε.Τ.: $\hat{\beta}(0) = \hat{\beta}$.

β) Μ ε ρ ο λ η ψ ί α: Από την προηγούμενη ιδιότητα έπεται ότι ο $\hat{\beta}(K)$ είναι ένας μεροληπτικός εκτιμητής.

$$E(\hat{\beta}(K) - \beta) = E(Z\hat{\beta} - \beta) = \ldots = (Z - I_p)\beta = \ldots = -K(X'X + KI_p)^{-1}\beta \neq 0$$

γ) Δ ι α κ ύ μ α ν σ η: $Var(\hat{\beta}(K)) = E[\{\hat{\beta}(K) - E(\hat{\beta}(K))\}\{\hat{\beta}(K) - E(\hat{\beta}(K))\}] =$

(10)

$$= E((Z\beta - Z\hat{\beta})(Z\beta - Z\hat{B})') = \sigma^2 Z(X'X)^{-1}Z' = \sigma^2(X'X + KIp)^{-1}(X'X)(X'X + KI_p)^{-1}$$

για $K = 0$ παίρνουμε: $Var(\hat{\beta}(0))$.

δ) Ε υ κ λ ε ί δ ι α ν ό ρ μ α: $\|\hat{\beta}(K)\| = (\hat{\beta}(K))' . (\hat{\beta}(K)) = \hat{\beta}' Z^2 \hat{\beta}$, που είναι μικρότερη εκείνης της Μ.Ε.Τ. Είναι δε συνεχής, μονότονος, αύξουσα συνάρτηση του Κ, τέτοια, ώστε όταν $K \longrightarrow \infty$ τότε $\|\hat{\beta}(K)\| \longrightarrow 0$. Επειδή $\|\hat{\beta}(K)\| < \|\hat{\beta}(0)\|$ ο Malinvaud (έκδ. 1983) προτείνει τον όρο «παλινδρόμηση (εκτιμητής) σμίκρυνσης» (raccourcie) και ο P. Cazes (βλ. (1)) τον όρο «περιορισμένος εκτιμητής» (borne). Είναι γεγονός ότι ο εκτιμητής Ridge βρίσκεται πιο κοντά στην αρχή των αξόνων αλλά το αποτέλεσμα σμίκρυνσης είναι γενικό και έχει ενδιαφέρον να εξετάζουμε καθεμιά συντεταγμένη του εκτιμητή (βλ. (4).

ε) Τ ο Μ. Τ. Σ. : $E(L^2) = E((\hat{\beta}(K)-\beta)'(\hat{\beta}(K)-\beta)) = E((Z\hat{\beta}-\beta)'(Z\hat{\beta}-\beta)=$

$=E((Z\hat{\beta}-Z\beta)'(Z\hat{\beta}-Z\beta)+(Z\beta-\beta)'(Z\beta-\beta))$, κατόπιν αντικαταστάσεως του $(-\beta)$

διά του $(-Z\beta+Z\beta-\beta). =\ldots= \sigma^2 tr((X'X)^{-1}Z'Z)+\beta'(Z-I)'(Z-I)\beta=$

$$(11) \qquad = \{\sigma^2 \sum_i \lambda_i/(\lambda_i+(K)^2\} + \{K^2\beta'(X'X+KI_p)^{-2}\beta\},$$

διά αντικαταστάσεως του $Z=I-K(X'X+KI)^{-1})=\gamma^1(K)+\gamma^2(K)=\Sigma.\Delta.+(\text{Με}\rho).^2$

ζ) Η Σ υ ν ο λ ι κ ή Δ ι α κ ύ μ α ν σ η $\gamma^1(K)$ είναι συνάρτηση συνεχής, μονότονος φθίνουσα του Κ, και το τετράγωνο της Μεροληψίας $\gamma^2 = (K)$ είναι συνάρτηση συνεχής, μονότονος, αύξουσα του Κ.

η) Θ ε ώ ρ η μ α ύ π α ρ ξ η ς τ ο υ Κ. *: Υπάρχει πάντοτε ένα $K>0$ τέτοιο ώστε :*

$E(L^2(K))<E(L^2(0)$ ή $\text{Μ.Τ.Σ.}.(\hat{\beta}(K))<\text{Μ. Τ. Σ.}.(\hat{\beta}(0)$.

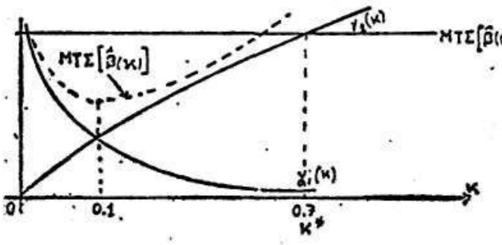
Αρκεί να δειχθεί ύτι :

$$\exists\ K > 0 : \frac{\theta\ E(L^2(K))\}}{\theta K} = \frac{\theta\gamma'(K)}{\theta K} + \frac{\theta\gamma^2(K)}{\theta K} < 0$$

(που σημαίνει ότι η συνάρτηση Μ.Τ.Σ..$(\widehat{\beta}(K))$ είναι φθίνουσα).Για τούτο αρκεί:

$$K < \sigma^2/\alpha^2_{max},\ \text{όπου}\ \alpha = P\beta,\ P'P = I.$$

Μια γραφική παράσταση του θεωρήματος δίνεται σχο σχήμα :



κ*: η άριστη του κ (optim

O Theobald (14) γενίκευσε το κριτήριο του Μ.Τ.Σ. σε :

$$M(K, B) = ((\hat{\beta}\ (K) - \hat{\beta})'\ B\ (\hat{\beta}\ (K) - \beta))$$

και όρισε έναν «καλό» εκτιμητή για K, τέτοιον ώστε :

$$M\ (K, B) - M\ (0, B) > 0$$

για κάθε θετικά ορισμένη μήτρα B. Χρησιμοποιώντας αυτόν τον ορισμό παρουσίασε την ικανή συνθήκη :

$$0 < K < 2\sigma^2/\beta'\beta,$$

και ο Obenchain το 1978 (12) παρουσίασε την ικανή και αναγκαία συνθήκη :

$$(0 < K < (2)/\ |l_p|),$$

όπου l_p είναι η αρνητική χαρακτηριστική ρίζα της $(X'X)^{-1} - \beta\beta'/\sigma^2$ (1).

2) Η Γενικευμένη Ridge (G.R.R.):

Η αρχή της G. R. R. συνίσταται στο να βρούμε μια σειρά από K_i τέτοια, ώστε :

$$\frac{\theta\,[\text{Μ.Τ.Σ.}\ (\alpha^*(\overline{K}))]}{\theta\overline{K}} = 0$$

όπου $\alpha^*(\overline{K})$ είναι η γενική μορφή της R.R που δίνεται από τη σχέση :

$$(12) \qquad \alpha^*(\overline{K}) = [(X^*)'\,(X^*) + \overline{K}]^{-1}\,(\chi^*)'Y = (\Lambda + \overline{K})^{-1}\,(X^*)'\,Y$$

με : $\overline{K}$ = διαγ. (K), μήτρα με μη μηδενικά τα διαγώνια στοιχεία. K_i, i=1, p.

με : $E\,(\alpha^*) = D\alpha$, D η διαγώνια μήτρα των $\lambda_i/(\lambda_i+K_i)$ i=1, ... p.

με : $\mathrm{Var}\,(\alpha^*) = \sigma^2\,(\Lambda + \overline{K})^{-1} \quad \Lambda(\Lambda+\overline{K})^{-1} = \sigma^2 \quad D^*$, D^* η διαγώνια μήτρα των τιμών $\lambda_i/(\lambda_i + K_i)^2$

* Οι ιδιότητες της G. R. R είναι παρόμοιες με εκείνες της O. R . R.

* Παρουσιάζεται το ίδιο πρόβλημα με την O.R.R., όσον αφορά την εκλογή της άριστης τιμής του K.

1. Δες και [11] για άλλες αλγεβρικές ιδιότητες της μεθόδου αυτής.

* Πάντα υπάρχει ένα $K > 0$ τέτοιο ώστε :

$$E(L^2(K)) < \dot{E}(L^2(0)).$$

$$* \quad \text{Μ.Τ.Σ.}(\alpha^*) = \sum_{i=1}^{p} \frac{\sigma^2\lambda_i + (K_i a_i^2}{(\lambda_i + K_i))^2}$$

όπου : $\alpha = P\beta$, $(X^*)'(X^*) = \Lambda$, $\alpha'\alpha = \beta'\beta$ και $P'P = I$.

3) Προσαρμοσμένοι εκτιμητές Ridge

Το σημαντικότερο μειονέκτημα της Ridge είναι η εκλογή της τιμής του K. Πρός τούτο, διάφοροι μέθοδοι έχουν προταθεί :

(α) Η Γραφική Μέθοδος (Ridge-Trace— R. T.) των Hoerl-Kennard και του Vinod: Η R.T, που προτάθηκε από τους συγγραφsείς Hoerl και Kennard παρουσιάζει την εξέλιξη των παραμέτρων σε συνάρτηση του KE(0,1). Μεγάλη μεταβλητικότητα των παραμέτρων για χαμηλές τιμές του K είναι δείκτης ύπαρξης της πολυσυγγραμμικότητας. Μόλις το σύστημα σταθεροεποιηθεί δεχόμαστε την αντιστοιχούσα τιμή του K (ή διάστημα των τιμών του K στο οποίο σταθεροποιείται το σύστημα). Η R.T. είναι μέθοδος υποκειμενική και εξαρτάται από την πείρα του σχεδιαστή. Όμοια είναι και η R.T. του Vinod (βλ. (15), με τη διαφορά ότι παρουσιάζουμε γραφικά τους εκτιμητές Ridge σε συνάρτηση του $m = p - \sum_i \lambda_i/(\lambda_i + K) = p - tr(X'X)(X'X + KI)^{-1}$. Το m ονομάζεται multicollinearity allowance.

(β) Μηχανικές Μέθοδοι: Σημείο εκκίνησης των Μ.Μ. είναι οι προσεγγίσεις των Hoerl - Kennard και Hoerl - Kennard - Baldwin :

$$K_H.K = \hat{\sigma}^2 / \sum_{1}^{p} \hat{\alpha}_i^2, \quad \text{όπου } \hat{\alpha}^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/(n - P).$$

ΚΑΝΟΝΕΣ	ΣΧΟΛΙΑ
$K_{H.K.} = \hat{\sigma}^2/\sum_i \hat{a}_i^2$	ΜΤΣ $(K_{H.K.}) <$ ΜΤΣ (Μ.Ε.Τ.)
$K_{H.K.B.} = p.\hat{\sigma}^2/\sum_i \hat{a}_i^2$	0 $K_{H.K.} \leq K_{H.K.B.} <$ Koptimal.
$K_{L.W.} = p.\hat{\sigma}^2/\sum_i \lambda_i \hat{a}_i^2$	Βασίζεται στην **Bayesian ανάλυση.** $K_{L.W.} > K_{H.K.}$
$\sum_i \hat{a}_i^2/\hat{\sigma}^2((1/K_D)+(1/\lambda_i)) = P$	Δύσχρηστος. Δίνει καλλίτερα αποτελέσματα από τους K_{WR}, K_{THD}, και K_{HKB}
$K_{THD} = (p-2).\ \hat{\sigma}^2/\hat{a}_i^2$	για $p>2$
$\hat{\sigma}^2.\ \sum_i \lambda_i/(\lambda i+K_{WR})^3 = K.\sum_i \dfrac{\lambda_i \hat{a}_i^2}{(\lambda_i+K)^3}$	Επιλύουμε ως προς K_{WR}
$\sum_i \hat{a}_i^2/((1/Ks)+(1/\lambda_i) = \rho.\hat{\sigma}^2((n-p)/(n-p-2))$	Φαίνεται ότι Ks και K_D είναι προτιμώτερα αλλά δύσχρηστα.
$K_{H.S.L.} = \hat{\sigma}^2 \sum_i \lambda_i^2 \hat{a}_i^2 / \sum_i \lambda_i^2 \hat{a}_i^4$	
$\sum_i \hat{a}_i^2/(\hat{\sigma}^2 K^{-1}_{D.S.W.} + \hat{\sigma}^2\lambda_i^{-1}) = P$	Βασίζεται στην **Bayesian** ανάλυση
$(\hat{\beta}(K))'(\hat{\beta}(K)) = \hat{\beta}'\hat{\beta} + \sigma^2\, tr\, (X'X)^{-1}$ όταν το K αυξάνει, τότε $(\hat{\beta}(K))'(\hat{\beta}(K))$ φθίνει.	Προτάθηκε από τους McDONALD και GALARNEAU το 1975. Τα αποτελέσματα δεν είσαφή ως προς την σύγκριση του $K_{D.G.}$ με τους άλλους κανόνες.
<u>HOERL - KENNARD 1981:</u> $(K_{t+1}' - K_t)/K_t \leq \Delta = 20.T^{1,3}$ (1) όπου $T = 1/p\ tr(X'X)^{-2}$	Βασίζεται στην G.R.R.. Αφού εκτιμήσουμε με την Μ.Ε.Τ.; $\frac{\hat{\sigma}^2}{\hat{a}_i^2}$ αντικαθιστούμε στο $\alpha^*(\bar{K})$. Επανεκτιμούμε $\hat{\sigma}^2/\hat{a}_i^2$, μετά την εκτίμηση του $\alpha^*(\bar{X})$ Επαναλαμβάνουμε τόσες φορές ως ότου σταθεροποιηθεί το σύστημα δηλαδή να ισχύει η (1).
<u>OBENCHAIN:</u> a) $b^*(k,q) = ((X'X)^{1-q}+KI)^{-1}(X'X)^{-q}X'Y$ b) $K_{OB} = \dfrac{(\sum_i \lambda_i^q)(\sum_j \lambda_j^{(1-q)})(1-R^2\cos^2(q))}{p.n.R^2.\cos^2(q)}$	α) Ορίζεται από $K_i = K.\lambda_i^q$. Όταν $q=0$ παίρνουμε O.R.R.. Όταν $q>0$ και μεγάλο, τότε $\lambda_i/(\lambda_i+K)$ μεγαλώνει. Όταν $q<0$ και μεγάλο, τότε $\lambda_i/(\lambda_i+K)$ παρόμοιο με την P.C.R β) (1981). Σε επίπεδο εφαρμογών φαίνεται πως δεν επιλύεται ((11)).

<u>Όπου:</u> K_{LW} = LAWLESS - WANG, K_D = DEMSTER, K_{THD} = THIMSTED (1976), Ks = SCOLVE (1973), K_{WR} = WERMUTH, K_{HSL} = HOCKING - SPEED - LYNN (1976), K_{DSW} = DEMSTER - SCHATROFF - WERMUTH (1977).

$$K_{\text{H.K.B}} = p\sigma^2 / \sum_i \alpha_i^2 .$$

Στον πίνακα της σελίδας 470 συλλέξαμε τις πιο γνωστές Μηχανικές Μεθόδους και τα αποτελέσματα που μέχρι τώρα έχουν δώσει.

III. ΔΙΑΦΟΡΕΣ ΕΡΜΗΝΕΙΕΣ ΤΗΣ RIDGE

(α) Εύκολα αποδεικνύεται ότι ο εκτιμητής $\hat{\beta}(K)$ μπορεί να γραφεί υπό τη μορφή: $\hat{\beta}(K) = (\tilde{X}'\tilde{X})^{-1}\tilde{\chi}'\,Y$ και να ερμηνευτεί σαν ένας εκτιμητής της Μ.Ε.Τ. με $\tilde{X} = X + \sum_i K_i/\sqrt{\lambda_i}\,P_i,\ Q_i$, όπου Q_i είναι μία διαγώνια μήτρα p διαστάσεων.

(β) Με τη μέθοδο των κυριωτέρων συνιστωσών παίρνουμε μία οικογένεια εκτιμητών από:

$$B_v^{P.C.} = ((X^*)'\,(X^*))_v^{-1}\,(X^*)',\ Y = \Lambda_v^{-1}\,(X^*)'Y, \text{ με } v<p$$

Είδαμε ότι $\alpha^*(\overline{K}) = (\Lambda + \overline{K})^{-1}\,(X^*)_,\ Y = D^{-1}\,(X^*)'Y$, όπου $D^{-1} =$ διαγ. $[(\lambda_i + K_i)^{-1}]$.

Αν λοιπόν μηδενίσουμε τις ν πρώτες συνιστώσες και κρατήσουμε τις (p—ν), τότε η G.R.R. ισοδυναμεί με την R.P.C.[1]

(γ) Μία οικογένεια εκτιμητών σύμφωνα με τη μέθοδο του Γενικευμένου Αντίστροφου Πίνακα (G.I.), δίνεται από τη σχέση:

---

1. Μπορούμε επίσης να αποδείξουμε την ισοδυναμία μεταξύ G. R. R. και της μεθόδου Stein-Like ή με την παλινδρόμηση υπό περιορισμούς.

$$\hat{\beta}^{+} = A_{\nu}^{+} X'Y, \text{ όπου } A_{\nu}^{+} = P_{\nu}\Lambda_{\nu}^{-1} P_{\nu}' = \sum_{i=1}^{\nu} \lambda_i^{-1} P_i P_i', \text{ με } \nu < p \text{ και}$$

$$A = XX' \qquad (9).$$

Αν προσθέσουμε ένα μικρό συντελεστή K στις χαρακτηριστικές ρίζες λ_i, τότε παίρνουμε την R.R.. Πρέπει όμως να σημειώσουμε ότι η R.R. δεν είναι G.I. γιατί δεν ικανοποιεί τη σχέση : $X'X\,(X'X)^{+}X'$, $X = X'X$.

(δ) Οι συγγραφείς Lindley και Smith (1972)) έδωσαν στην R.R.μία ε ρ μ η - ν ε ί α B a y e s i a n : Αν η κατανομή a priori του β : $\beta \sim N(0, \omega,^2 I_p)$, τότε η κατανομή a posteriori είναι : $\beta \sim N[\{(X'X + (\sigma^2/\omega^2)I)^{-1}X'Y\}, \sigma^2\{X'X + (\sigma^2/\omega^2)I\}^{-1}]$ όπου ο εκτιμητής Ridge $K = \sigma^2/\omega^2$ παρουσιάζεται σαν η μαθηματική ελπίδα της κατανομής α posteriori της παραμέτρου β η μαθηματική ελπίδα της κατανομής α priori της β είναι ίση με μηδέν).

IV. ΕΠΕΚΤΑΣΗ ΤΗΣ R.R. ΣΤΑ ΣΥΣΤΗΜΑΤΑ ΑΛΛΗΛΟΕΞΑΡΤΩΜΕΝΩΝ ΕΞΙΣΩΣΕΩΝ

Ας πάρουμε τη διαρθρωτική μορφή :

$$(13) \qquad Y_{\beta} + X_{\gamma} + U = 0$$

ή

$$Y_1 + \beta_{12} Y_2 + \ldots + \beta_{1p} Y_p + \gamma_{11} X_1 + \ldots + \gamma_{1q} X_q + U_1 = 0$$

$$\beta_{p1} Y_1 + \beta_{p2} Y_2 + \ldots + Y_p + \gamma_{p1} X_1 + \ldots + \gamma_{pq} X_q + U_2 = 0.$$

(E) Εκ κατασκευής του υποδείγματος η υπόθεση ότι δεν υπάρχουν ακριβείς γραμμικές σχέσεις μεταξύ των ερμηνευτικών μεταβλητών παραβιάζεται και έτσι δε μπορούμε να εκτιμήσουμε το υπόδειγμα με την Μ.Ε..Τ.

1) Η Μ. Ε. Τ. σε δύο στάδια:

Ας γράψουμε την (13) υπό τη μορφή:

(14) $\underline{Y} = Y_0\beta_0 + X_0\gamma_0 + U$

όπου,

$\underline{Y}X = (Y_0, X_0)\,\delta + U$, με $\delta = \binom{\beta_0}{\gamma_0}$. $\underline{Y}$ είναι η στήλη του Y με n στοιχεία.

Y_0 (αντιστ. X_0) είναι η μήτρα np_1 των παρατηρήσεων p_1 των ενδογενών μεταβλητών (αντιστ. nq_1 των q_1 εξωγενών μεταβλητών.). Ως γνωστό, η Μ.Ε.Τ. σε δύο στάδια είναι ισοδύναμη με τη μέθοδο των βοηθητικών μεταβλητών (I.V.). Οι υπολογισμένες τιμές των ενδογενών μεταβλητών μπορούν να χρησιμεύουν σαν I.V. γιατί δε συσχετίζονται με το διαταρακτικό όρο, ενώ συσχετίζονται με τις πραγματικές τιμές. Θα αντικαταστήσουμε δηλαδή την Y με μία I.V. για την οποία η ανωτέρω υπόθεση (Η) δε παραβιάζεται.

Για το λόγο αυτό διαλέγουμε την $\hat{Y}_o$ που παίρνουμε από την Μ.Ε.Τ. εκτιμώντας κάθε $Y_k \in Y_o$ σε συνάρτηση του συνόλου των q εξωγενών μεταβλητών.

Πρώτο Στάδιο: Αντικαθιστώντας στή (14):

$$\forall Y_K, \in Y_o; \qquad \hat{Y}_K = \sum_{i=1}^{q} \Phi_i,\ X_i + V, \quad \Phi_i = (\Phi_1, \ldots, \Phi_q) \text{ το διάνυσμα}$$

των προς εκτίμηση παραμέτρων, $X_i = (X_1 \ldots, \chi_q)$ το διάνυσμα των εξωγενών μεταβλητών. Με την Μ.Ε.Τ. παίρνουμε :

$$\hat{\Phi} = (X'X)^{-1}\,X'Y_K \ \text{ και } \ \hat{Y}_K = X\hat{\Phi} = X\,(X'X)^{-1}\,X'Y_K.$$

Δ ε ύ τ ε ρ ο Σ τ ά δ ι ο: Αντικαθιστώντας το Y_o με $\hat{Y}_o$ η (14) γράφεται:

$$Y = (\hat{Y}_o, X_o)\,\delta + U \quad ή \quad Y = Z\delta + U,$$

όπου $Z = (\hat{Y}_o X_{o,o})$ (οι μεταβλητές $\hat{Y}_κ$, Y_o και Z λέγονται «βοηθητικές»). Εφαρμόζοντας για δεύτερη φορά την Μ.Ε.Τ. παίρνουμε:

$$(15) \quad D = \hat{\delta} = (Z'Z)^{-1} Z'Y \quad ή$$

$$\hat{\delta} = \begin{pmatrix} Y_o' Y_o - V_o' V_o & Y'_o X_o \\ X_o Y_o & X_o' X_o \end{pmatrix}^{-1} \begin{pmatrix} Y_o' - V_o' \\ X_o' \end{pmatrix} Y$$

όπου $V_o = Y_o - X\,(X'X)^{-1}\,X'Y_o$.

Σε κάθε στάδιο οι υποθέσεις του γραμμικού παραδείγματος είναι απαραίτητες. Αριθμός μεγαλύτερος του αριθμού των εξωγενών μεταβλητών αυξάνει την πιθανότητα εμφάνισης της πολυσυγγραμμικότητας στο πρώτο στάδιο πάνω στο χ, και στο δεύτερο στάδιο πάνω στο Ζ. Μπορούμε λοιπόν να προτείνουμε σαν εναλλακτική λύση, τη μέθοδο R.R σε δύο στάδια (για υποδείγματα μικρών διαστάσεων).

2) Η R. R. σ ε δ ύ ο σ τ ά δ ι α (D. R. R.):

Στο π ρ ώ τ ο σ τ ά δ ι ο αντί να μετασχηματίσουμε το $\hat{Y}_o$ σε Y_o με τη Μ.Ε.Τ. μπορούμε να χρησιμοποιήσουμε την R.R.O.:

$$(16) \qquad \hat{Y}_o = X\hat{\Phi}\,(K_1) = X\,(X'X + K_1 I)^{-1}\,X'Y_o$$

Σ τ ο δ ε ύ τ ε ρ ο σ τ ά δ ι ο παίρνουμε: με $Y = Z\delta + U$

$$(17) \qquad \hat{\delta}\,(K_2) = (Z'Z + K_2 I)^{-1}\,Z'Y$$

Ομοίως, έχουμε την D. G. R. R. Στο πρώτο στάδιο $\hat{Y}_o = X\hat{\Phi}(K_1) =$

$= X (X'X + P'K_1P)^{-1} X'Y_o$ και στο δεύτερο στάδιο $\hat{\delta}(K_2) =$

$= (Z'Z + P' K_2P)^{-1} Z'Y$, με $P'P = I$ και P η μήτρα των χαρακτηριστικών διανυσμάτων της Z'Z. Πρόβλημα πάντως παραμένει η εκλογή των άριστων τιμών των συντελεστών K_1, K_2. (1)

V. ΜΙΑ ΕΦΑΡΜΟΓΗ ΤΗΣ R. R. O.

Σε αυτήν την παράγραφο εκτιμήσαμε ένα υπόδειγμα με την R.R. Πρόκειται για το υπόδειγμα .·

$$I_t = \alpha + \beta_1 Y_t + \beta_2 Y_{t-1} + \beta_3 (ITC)_t + U_t,$$

όπου :

It οι επενδύσεις της περιόδου t, Yt το Εθνικό εισόδημα της ίδιας περιόδου και ITC (Investment Tax Credit) ένα φορολογικό κίνητρο επενδύσεων που θεσπίστηκε το 1962, ανεστάλη το 1966 και επαναφέρθηκε το 1970 στις Η.Π.Α. Στο υπόδειγμα μας I.T.C. είναι μία ψευδομεταβλητή με : I.T.C. =0 για τα έτη που δεν υπήρχε το κίνητρο. Ο συντελεστής β3 είναι ένα μέτρο της αποτελεσματικότητας του κινήτρου ($\beta_3 > 0$). Το υπόδειγμα καθώς και τα δεδομένα τα πήραμε από το βιβλίο του Intriligator σελ. 83 - 84. Σημειώνουμε ότι δεν θα ασχοληθούμε πολύ με τη οικονομική ερμηνεία των αποτελεσμάτων.

Η μέθοδος Μ. Ε. Τ.(2) μας έδωσε τα εξής αποτελέσματα :

$$\hat{I}_t = 0.0447 + 1,365Y_t - 0,364Y_{t-1} - 0,0762 \text{ I.T.C.}$$

---

(1). Στο επίπεδο αυτό, ο J: P. Marciano (9) παρουσίασε πρόσφατα μια στατιστική υπολογιστική διαδικασία, πολύπλοκη μεν αλλα αποτελεσματική.

(2). Τα δεδομένα τα μετασχηματίσαμε σε : $(\sqrt{n}\,\overline{S_1})^{-1}(X_1 - \overline{X_1})$ («Standard Form») [10].

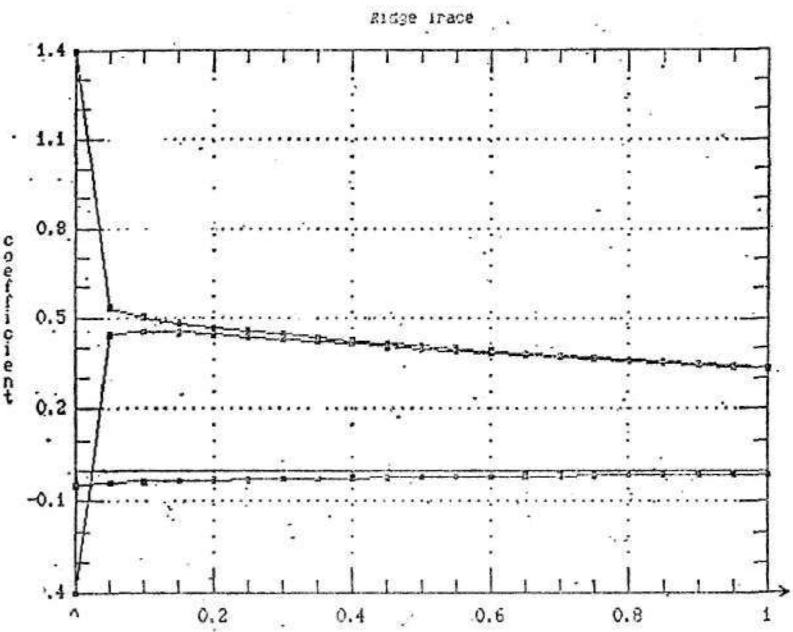
τυπικά σφάλμ.	(0,0185)	(0,188))	(0,188)	(0,0242).
t του Student	(2.422)	(7,262)	(—1,94)	(—3,146)

$R^2 = 0,9988$ και ο συντελεστής συσχετίσεως $Y_tY_{t-1}=0,9978$

Η τιμή του t από τους πίνακες για $\alpha = 5\%$ είναι 2,306 απ' όπου έπεται η μή σπασημαντικότατα του συντελεστή $\hat{\beta}_2$, Παρατηρούμε ακόμα ότι : (α) αλλαγή προσήμου μου του $\hat{\beta}_2$ για $K > 0.025$, (β) $E(L^2)=42,115\sigma^2$ που σημαίνει ότι η μέση απόσταση μεταξύ $\hat{\beta}$ (της Μ.Ε.Τ. και της παραμέτρου β είναι 42 φορές περίπου μεγαλύτερη εκείνης όταν οι μεταβλητές είναι ορθογώνιες. Ας σημειώσουμε ακόμα ότι και οι έλεγχοι κατά Farar - Glauber και κατά Haitovsky, που εφαρμόσαμε, συμφωνούν στην ύπαρξη πολυσυγγραμμικότητας. Πιστεύουμελοιπόν ότι οι εκτιμήσεις των συντελεστών με την Μ.Ε.Τ. δεν είναι ακριβείς.Ετσι,εφαρμόσαμε την παλινδρόμηση Ridge δίνοντας στον συντελεστή Κ διάφορες τιμές, από τις οποίες παρουσιάζουμε τις δέκα : $K \in (0, 0,,1,0..0,2.....,1)$ στο παρακάτω σχήμα. (1).

Από το ίδιο σχήμα που παρουσιάζεται η R.T. των εκτιμητών μπορούμε να πούμε ότι το σύστημα σταθεροποιείται για τιμές του $K \in (0,05$ -0,2). (Επιπρόσθετη ένδειξη αποτελεί η υψηλή τιμή του R^2). Από το ίδιο σχήμα παρατηρούμε ότι οι εκτιμήσεις με τη Μ.Ε.Τ. για Y_t και Y_{t-1} ITC είναι υπερεκτιμημένες για την πρώτη και υποεκτιμημένες για τις δεύτερες. Στην πρώτη περίπτωση γιατί β_1 φθίνει όσο το Κ αυξάνει, και στη δεύτερη περίπτωση γιατί β_2 και β_3 αυξάνουν όσο το Κ αυξάνει.

1. Δεν παρουσιάζουμε τα γραφήματα : Το Άθροισμα των Τετραγώνων των Καταλοίπων και την Τετραγωνική Απόσταση των Εκτιμήσεων Ridge. Η κλίση των καμπυλών αυτών συμφωνεί με τη θεωρία. Ας σημειώσουμε μόνο ότι το Α.Τ.Κ.. είναι μικρότερο σε σχέση με το αντίστοιχο της Μ.Ε.Τ. και ότι η Τ.Α.Ε. Ridge [βι (K)' βι (K)] δίνεται υψηλότερη ύστερα από την αλλαγή πρόσημου του β2 (κριτήριο εκλογής του Κ κατά Brown - Beattie).



Μπορούμε ολοιπόν, για παράδειγμα, να διαλέξουμε την «λογική» τιμή $K = 0.15$ εντός του διαστήματος (0.05 - 02) για την οποία το σύστημα παρουσιάζει μια σταθερότητα και για την οποία παίρνουμε : $\beta_0(K)=0.03$ $\beta_1(K) = 0.498$ $\beta_2(K) = -0.451$ και $\beta_3(K) = -0.034$ για $K = 0.15$ [1].

Εφαρμόσαμε στη συνέχεια τρεις από τις Μηχανικές Μεθόδους ($K_{H \cdot K \cdot}$, $K_{H \cdot K \cdot B \cdot}$ $K_{T \cdot H \cdot D \cdot}$) και οι τιμές του K που πήραμε βρίσκονται εντός του διαστήματος (0,000811 - 0,023005) που βρίσκονται πολύ κοντά στο 0 και που είναι κατά πολύ μικρότερες από εκείνες της R. T. Από πρακτικής μάλιστα πλευράς έχει αποδειχθεί ότι η μικρή τιμή του K ($K<0.01$) απαιτείται μόνο όταν το υπόδειγμα δεν περιέχει σταθερό όρο (βλ. (10). Η R.T. όμως είναι καλύτερη μέθοδος από τιςΜΜ. για την εκλογή της τιμής του K γιατί δίνει μεγαλύτερες τιμές και γιατί αν αντικαταστήσουμε τις τιμές των Μ Μ. στο $E(L^2)$ παρατηρούμε ότι η ευστάθεια των εκτιμητών δεν έχει ακόμα πραγματοποιηθεί.Τέλος, δεν λαμβάνουν υπόψιν τυχόν αλλαγές προσήμων των εκτιμητών : στο συγκεκριμένο μάλιστα παράδειγμα για τιμές εντός του διαστήματος των Μ.Μ. ο β_2 παραμένει αρνητικός.

1. Σημειώνουμε ότι οι εκτιμητές είναι στατιστικά σημαντικοί σε επίπεδο $\alpha = 5\,\%$ και $R^2=0.97$ Όσον αφορά τον β_3, το μόνο ίσως που μπορούμε να πούμε, είναι ότι το φορολογικό αυτό κίνητρο δεν εξυπηρετεί τον σκοπό της θεσπίσεώς του.

Σ ΥΜΠΕΡΑΣΜΑ

Παρουσία του προβλήματος της πολυσυγγραμμικότητας σε ένα υπόδειγμα η Μ.Ε.Τ. δίνει αποτελέσματα ασταθή και αναξιόπιστα. Ανάμεσα στις διάφορες μεθόδους επίλυσης του προβλήματος αυτού είναι η R.R. που εν συντομία παρουσιάσαμε. Η ανάλυση Ridge πρωτοπαρουσιάστηκε από τον Hoerl αλλά η Ridge Regression ολοκληρώθηκε σε δύο άρθρα των Hoerl - Kennard το 1970.

Το πλεονέκτημα της μεθόδου αυτής είναι ότι μας επιτρέπει να ανακαλύψουμε τη δομή των μεταβλητών, να αναλύσουμε την ευαισθησία των εκτιμώμενων συντελεστών των παραμέτρων και να βρούμε ένα σύνολο από εκτιμητές πιο κοντά στις πραγματικές τιμές των παραμέτρων σύμφωνα με το στατιστικό κριτήριο του Μ. Τ.Σ. Αξίζει επίσης να προσθέσουμε την ευκολία των υπολογισμών. Για την εκλογή της τιμής του Κ, δέκα έως τριάντα αντιστροφές της μήτρας $(X'X+KI\rho)$ αρκούν για να μπορέσει ο σχεδιαστής να προσδιορίσει πού η R.T. σταθεροποιείται. Το μειονέκτημα της είναι η εκλογή της άριστης τιμής του Κ και τούτο γιατί η εκλογή αυτή είναι υποκειμενική. Το μειονέκτημα αυτό έκανε την R. R. λιγότερο δημοφιλή από τις άλλες μεθόδους για τούτο απαιτείται η περισσότερη θεωρητική της ανάπτυξη συνοδευμένη από εφαρμογές.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Cazes P. : «Méthodes de Regression, I - III,» Les Cahiers de Γ Analyse des données, 1978, Vol. m, No 2.

2. ΓΚΑΜΑΛΕΤΣΟΣ Θ.: «Θεωρητική Οικονομετρία I - II», 1981.

3. EGERTON M.: — LAY COCK P. : «Some Criticisms of Stochastic Schrinkage and R. R, with Countrexamples», Technovetrics, 23 (2) , 1981.

4. FOURGEAUD - GOURIEROUX - PRADEL : «La Regression Ridge», Cahiers du Séminaire d'Econometrie, 25, 1982.

5. GIRAUD R. : «Le Problème des presques colinearites», In Cahiers d' Econometrie Appliquee, B. MUNIER - M. EGEA eds., 1985, Univ. Aix - Marseille III France.

6. HOERL A. E. : - KENNARD R. W. : «Ridge Regression : Biased Estimation for Nonorthogonal Problems», και «Ridge Regression : Applications to Monorthogonal Problems» Technometrics, 12 (1) 1970.

7. IMTRILIGATOR M. : «Οικονομετρικά Υποδείγματα I- II», Gutenberg.

8. LAZARIDIS A. : «A note regarding the problem of perfect multicolinearity», Indian Journal of Economics και «A Straight - toward approach to the problem of multicolinearity in econometrics», Spoudai, A. 1, 1980.

9. MARCIANO J. P. - MEGEA - P. GUERINI. Dèductibilitè et langages dans les Systèmes Experts : application à la théorie de Γ estimation». Symposium International of Forecasting, Mondreal Canada 1985.

10. MARQUARDTD. : «Generalized Inverse, Ridge Regression, Biased Linear Estimation and Non - linear Estimations», Technometrics, 12 (3), 1970 και MARQUARDT D. - SNEE R. : «Ridge Regression in practice», The American Statistician, 29 (1), 1975.

11. McDONALDC: «Some Algebraic Properties of Ridge Coefficients», J.R.S., S. B. 42(1), 1980.

12. OBENCHAIN R. : «Good and Optimum Ridge Estimators», Ann. Stat., 6, 1978.

13. ΠΑΝΑΣ Ε. : «Το πρόβλημα της πολυσυγγραμμικότητας : Παρατηρήσεις», Σπουδαί ΚΖ, 3, 1977.

14. THEOBALD C. : «Generalisation of M.S.E. applied to Ridge Regression», J.R.S.S. B. 36, 1974.

15. VINOD H. - ULLAH A.: «Recent Advances in Regression Methods» 1981.