Econometric Theory, **18**, 2002, 1461–1465. Printed in the United States of America.
DOI: 10.1017/S0266466602186099

PROBLEMS AND SOLUTIONS

PROBLEMS

02.6.1. *Oblique Projectors*, proposed by Götz Trenkler. Let $\mathbf{P}$ be an idempotent matrix with possibly complex entries. Assume that $\mathbf{P}$ is not Hermitian; i.e., $\mathbf{P}$ differs from its conjugate transpose. Show that the Moore–Penrose inverse $\mathbf{P}^+$ is not idempotent.

02.6.2. *Autoregression and Redundant Instruments*, proposed by Stanislav Anatolyev. Consider a zero mean stationary autoregressive model of order k with independent and identically distributed (i.i.d.) innovations having variance σ^2:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_k y_{t-k} + \varepsilon_t.$$

It is well known that the efficient generalized method of moments (GMM) estimator of $\rho = (\rho_1\ \rho_2\ \ldots\ \rho_k)'$ based on the instrumental vector $z_t = (y_{t-1}\ y_{t-2}\ \ldots\ y_{t-k}\ y_{t-k-1}\ \ldots\ y_{t-\ell})'$ consisting of the last $\ell > k$ lags of y_t effectively exploits information in the most recent k lags of y_t (see, e.g., Kim, Qian, and Schmidt, 1999). In other words, the instruments $y_{t-k-1},\ldots,y_{t-\ell}$ are redundant (see Breusch, Qian, Schmidt, and Wyhowski, 1999) given $y_{t-1},\ldots,y_{t-k}$.

Prove the following more general proposition: when one uses the instrumental vector $z_t = (y_{t-p}\ y_{t-p-1}\ \ldots\ y_{t-p-k+1}\ y_{t-p-k}\ \ldots\ y_{t-p-\ell+1})'$ for $p \geq 1$, the instruments $y_{t-p-k},\ldots,y_{t-p-\ell+1}$ are redundant given $y_{t-p},\ldots,y_{t-p-k+1}$.

REFERENCES

Breusch, T., H. Qian, P. Schmidt, & D. Wyhowski (1999) Redundancy of moment conditions. *Journal of Econometrics* 91, 89–111.
Kim, Y., H. Qian, & P. Schmidt (1999) Efficient GMM and MD estimation of autoregressive models. *Economics Letters* 62, 265–270.

SOLUTIONS

01.6.1. *Minimax Median*—Solution,[1] proposed by Geert Dhaene.

(a) The cumulative distribution function (c.d.f.) of $\hat\beta$ is

$$F_{\hat\beta}(x) = \Pr\Big[\min_j\Big\{\max_i\{X_{ij}\}\Big\} \le x\Big]$$
$$= 1 - [1 - (F(x))^{r_n}]^{c_n},$$

 0266-4666/02 $9.50

using well-known properties of order statistics. Median-unbiasedness of $\hat{\beta}$ requires $F_{\hat{\beta}}(\beta) = 0.5$, and therefore

$$(1 - 0.5^{r_n})^{c_n} = 0.5. \tag{1}$$

Clearly, this implies $r_n \to \infty$ and $c_n \to \infty$ as $n \to \infty$. Hence we can write

$$(1 - 0.5^{r_n})^{c_n} = \exp(-0.5^{r_n} c_n) + o(1) = 0.5.$$

Solving for r_n yields

$$r_n = \frac{\log(c_n)}{\log(2)}(1 + o(1)). \tag{2}$$

Furthermore,

$$\begin{aligned}\log(c_n) &= \log(n) - \log(r_n)\\ &= \log(n) - \log(\log(c_n)) - \log(\log(2)) + o(1)\\ &= \log(n)(1 + o(1)),\end{aligned} \tag{3}$$

because

$$\lim_{n\to\infty} \frac{\log(\log(c_n))}{\log(n)} = \lim_{c_n\to\infty} \frac{\log(\log(c_n))}{\log(c_n)} = 0.$$

Combining (2) and (3) yields

$$r_n = \frac{\log(n)}{\log(2)}(1 + o(1)) = \log_2(n)(1 + o(1)).$$

(b) It will be shown that $r_n(\hat{\beta} - \beta)$ converges in distribution, provided that F is differentiable in an open neighborhood of β and has a positive derivative at β. Start from

$$\begin{aligned}\Pr[r_n(\hat{\beta} - \beta) \le x] &= 1 - [1 - [F(r_n^{-1}x + \beta)]^{r_n}]^{c_n}\\ &= 1 - \exp[-[F(r_n^{-1}x + \beta)]^{r_n} c_n] + o(1).\end{aligned} \tag{4}$$

Then, for some a_n in $[\beta, r_n^{-1}x + \beta]$ and sufficiently large n,

$$\begin{aligned}[F(r_n^{-1}x + \beta)]^{r_n} c_n &= [F(\beta) + r_n^{-1} x f(a_n)]^{r_n} c_n\\ &= 0.5^{r_n} c_n [1 + 2r_n^{-1} x f(a_n)]^{r_n}.\end{aligned} \tag{5}$$

From (1), $c_n = -\log(2)/\log(1 - 0.5^{r_n})$, wherefrom

$$\begin{aligned}\lim_{n\to\infty} 0.5^{r_n} c_n &= -\log(2) \lim_{z\to 0} \frac{z}{\log(1 - z)}\\ &= \log(2).\end{aligned} \tag{6}$$

https://doi.org/10.1017/S0266466602006096 Published online by Cambridge University Press

Moreover,

$$\lim_{n\to\infty} [1 + 2r_n^{-1} x f(a_n)]^{r_n} = \exp(2xf(\beta)). \quad \textbf{(7)}$$

Combining (4)–(7) gives the limiting c.d.f. of $r_n(\hat{\beta} - \beta)$:

$$\lim_{n\to\infty} \Pr[r_n(\hat{\beta} - \beta) \le x] = 1 - 2^{-\exp(2xf(\beta))}.$$

NOTE

1. An excellent solution has been independently proposed by S. Portnoy and G.W. Bassett, the authors of the problem. They noted that the limit distribution belongs to the class of Gompertz distributions or shifted and rescaled type III extreme value distributions. G. Dhaene also pointed out a typo in the statement of the problem: in part (b) of the problem the "/" should be omitted from the first sentence.

01.6.2. *Identification of Parameters in Two Competing Risk Models*—Solution,[1] proposed by S.K. Sapra.

Model 1. The crude hazard rates (in the presence of both risks) and the integrated hazard rates corresponding to the joint survival function in (1) are, respectively,

$$h_j(t) = -\partial \ln S(t_1, t_2)/\partial t_j|_{t_1=t_2=t} = \alpha\lambda_j[(\lambda_1 + \lambda_2)t]^{\alpha-1}, \qquad j = 1,2, \quad \textbf{(3)}$$

$$H_j(t) = \int_0^t h_j(u)du = \lambda_j(\lambda_1 + \lambda_2)^{\alpha-1}t^\alpha, \qquad j = 1,2. \quad \textbf{(4)}$$

Case (a). The likelihood function is (Kalbfleisch and Prentice, 1980, p. 181)

$$L = \prod_{i=1}^{n} (h_1(t_i))^{\delta_i}(h_2(t_i))^{1-\delta_i} \exp(-H_1(t_i) - H_2(t_i)), \quad \textbf{(5)}$$

and thus

$$\ln L = \sum_{i=1}^{n} \ln\alpha + \delta_i \ln\lambda_1 + (1-\delta_i)\ln\lambda_2 + (\alpha-1)\ln(\lambda_1+\lambda_2)t_i \\ - (\lambda_1+\lambda_2)^\alpha t_i^\alpha. \quad \textbf{(6)}$$

Because the information matrix, $I(\alpha, \lambda_1, \lambda_2)$, is nonsingular, all of the parameters α, λ_1, λ_2 are locally identified (Rothenberg, 1971).

Case (b). For $T = \min(T_1, T_2)$, the overall survival function is

$$S_T(t) = P(T_1 > t, T_2 > t) = S(t,t) = \exp(-(\lambda_1+\lambda_2)^\alpha t^\alpha). \quad \textbf{(7)}$$

https://doi.org/10.1017/S0266466602006096 Published online by Cambridge University Press

The likelihood function based on T alone is

$$L^* = \prod_{i=1}^{n} [\alpha(\lambda_1 + \lambda_2)^\alpha t_i^{\alpha-1}] \exp(-(\lambda_1 + \lambda_2)^\alpha t_i^\alpha), \tag{8}$$

and thus

$$\ln L^* = \sum_{i=1}^{n} \ln \alpha + \alpha \ln(\lambda_1 + \lambda_2) + (\alpha - 1) \ln t_i - (\lambda_1 + \lambda_2)^\alpha t_i^\alpha. \tag{9}$$

Because the information matrix, $I(\alpha, \lambda_1 + \lambda_2)$, is nonsingular, α and $\lambda_1 + \lambda_2$ are locally identified. However, λ_1 and λ_2 are not identified because the information matrix, $I(\alpha, \lambda_1, \lambda_2)$, is singular.

Indeed, the first-order conditions for maximization of $\ln L^*$ yield only two distinct equations for three parameters because the normal equations for λ_1 and λ_2 are identical:

$$\partial \ln L^*/\partial\lambda_1 = \partial \ln L^*/\partial\lambda_2 = n\alpha/(\lambda_1 + \lambda_2) - \alpha \sum_{i=1}^{n} (\lambda_1 + \lambda_2)^{\alpha-1} t_i^\alpha = 0. \tag{10}$$

Model 2. The crude hazard rates corresponding to the joint survival function in equation (2) are

$$h_j(t) = -\partial \ln S(t_1, t_2)/\partial t_j|_{t_1=t_2=t} = (\alpha + 1)\lambda_j/[(\lambda_1 + \lambda_2)t + 1], \qquad j = 1,2, \tag{11}$$

and the integrated hazard rates are

$$H_j(t) = [(\alpha + 1)\lambda_j/(\lambda_1 + \lambda_2)] \ln[(\lambda_1 + \lambda_2)t + 1], \qquad j = 1,2. \tag{12}$$

Case (a). The likelihood function (from (5), (11), and (12)) is

$$L = \prod_{i=1}^{n} \lambda_1^{\delta_i} \lambda_2^{1-\delta_i} (\alpha + 1)/[(\lambda_1 + \lambda_2)t_i + 1]^{\alpha+2}, \tag{13}$$

and thus

$$\ln L = \sum_{i=1}^{n} \delta_i \ln \lambda_1 + (1 - \delta_i) \ln \lambda_2 + \ln(\alpha + 1) - (\alpha + 2) \ln[(\lambda_1 + \lambda_2)t_i + 1]. \tag{14}$$

Because the information matrix, $I(\alpha, \lambda_1, \lambda_2)$, is nonsingular, all of the parameters α, λ_1, λ_2 are locally identified.

Case (b). For $T = \min(T_1, T_2)$, the overall survival function is

$$S_T(t) = P(T_1 > t, T_2 > t) = S(t, t) = 1/[(\lambda_1 + \lambda_2)t + 1]^{\alpha+1}. \tag{15}$$

https://doi.org/10.1017/S0266466602006096 Published online by Cambridge University Press

Therefore, the likelihood function based on T alone is

$$L^* = \prod_{i=1}^{n} (\alpha + 1)(\lambda_1 + \lambda_2)/[(\lambda_1 + \lambda_2)t_i + 1]^{\alpha+2}, \tag{16}$$

and the log likelihood function is

$$\ln L^* = \sum_{i=1}^{n} \ln(\alpha + 1) + \ln(\lambda_1 + \lambda_2) - (\alpha + 2)\ln[(\lambda_1 + \lambda_2)t_i + 1]. \tag{17}$$

Because the information matrix, $I(\alpha, \lambda_1 + \lambda_2)$, is nonsingular, α and $\lambda_1 + \lambda_2$ are locally identified. However, λ_1 and λ_2 are not identified because the information matrix, $I(\alpha, \lambda_1, \lambda_2)$, is singular.

Indeed, the first-order conditions for maximization of $\ln L^*$ yield only two distinct equations for three parameters because the normal equations for λ_1 and λ_2 are identical:

$$\begin{aligned}\partial \ln L^*/\partial\lambda_1 = \partial \ln L^*/\partial\lambda_2 &= n/(\lambda_1 + \lambda_2) - (\alpha + 2)\sum_{i=1}^{n} [t_i/[(\lambda_1 + \lambda_2)t_i + 1]] \\ &= 0. \end{aligned} \tag{18}$$

NOTE

1. The author thanks a referee for insightful comments.

REFERENCES

Kalbfleisch, J.D. & R.L. Prentice (1980) *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Rothenberg, T. (1971) Identification in parametric models. *Econometrica* 39, 577–591.

https://doi.org/10.1017/S0266466602006096 Published online by Cambridge University Press