Econometric Theory, **19**, 2003, 691±705. Printed in the United States of America.
DOI: 10.1017/S0266466603004122

PROBLEMS AND SOLUTIONS

02.3.2. *Badly Weighted Least Squares*, proposed by R. Koenker and S. Portnoy.

Consider the classical linear regression model

$$y_i = x_i'\beta + u_i$$

with $V(u_i) = \sigma_i^2 > 0$, u_i independent over $i = 1,2,\dots,n$, and $X = (x_i')_{i=1}^n$ of full column rank p.

Let $\Omega = \operatorname{diag}(\sigma_1,\dots,\sigma_n)$ and assume $\sigma_i \neq \sigma_j$ for some $i \neq j$. The asymptotic covariance matrix of the weighted least squares estimator

$$\hat\beta(a) = (X'\Omega^{-2a}X)^{-1}X'\Omega^{-2a}y$$

is the Eicker–White matrix, $(X'X)^{-1}X'\Omega^2X(X'X)^{-1}$ for $a = 0$, whereas for $a = 1$ we have $(X'\Omega^{-2}X)^{-1}$. For $a = \frac{1}{2}$ we obtain an expression for the asymptotic covariance matrix of the median regression estimator.

This *suggests* that in particular,

$$(X'X)^{-1}X'\Omega^2X(X'X)^{-1} \geq (X'\Omega^{-1}X)^{-1}X'X(X'\Omega^{-1}X)^{-1}$$

and more generally that, for any $\delta \in \mathbb{R}^p$,

$$e_{a,\delta} = \delta'(X'\Omega^{-2a}X)^{-1}X'\Omega^{2-4a}X(X'\Omega^{-2a}X)^{-1}\delta$$

is monotone in $a \in [0,1]$. Prove or disprove the monotonicity claim.

SOLUTIONS

02.3.2. Badly Weighted Least SquaresÐ Solution[1]

Douglas P. Wiens
University of Alberta, Edmonton, Alberta

Let $\mathbf{X}_{n\times p}$ have full column rank $p < n$ and let $\boldsymbol{\Omega}$ be a diagonal matrix whose diagonal elements $\{\omega_i\}_{i=1}^n$ are positive and not all equal. The discussion by Professors Koenker and Portnoy suggested that, for any (nonzero) $\boldsymbol{\delta} \in \mathbb{R}^p$, the function

$$e_{a,\boldsymbol{\delta}} = \boldsymbol{\delta}'(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}^{2-4a}\mathbf{X}(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1}\boldsymbol{\delta}$$

satisfies $e_{0,\boldsymbol{\delta}} \geq e_{0.5,\boldsymbol{\delta}}$ and more generally that $e_{a,\boldsymbol{\delta}}$ is a decreasing function of $a \in [0,1]$.

We shall show that the second suggestion, and hence the first also, is true for $p = 1$. Both can fail if $p > 1$; we illustrate this by an example in which $e_{a,\boldsymbol{\delta}}$ is not only nonmonotone but is *maximized* at $a = 0.5$.

First let $p = 1$, so that $\mathbf{X} = \mathbf{x}$, a column vector. We calculate that

$$\frac{d}{da} e_{a,\delta} = \frac{-2\delta^2 M}{\mathbf{x}'\boldsymbol{\Omega}^{-2a}\mathbf{x}}, \tag{1}$$

where

$$M = 2\mathbf{v}'\boldsymbol{\Omega}^{2-2a}\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{v} - 2(\mathbf{v}'\boldsymbol{\Omega}^{2-2a}\mathbf{v})(\mathbf{v}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{v}),$$

$\mathbf{L}_{\boldsymbol{\Omega}}$ is the diagonal matrix with diagonal elements $\{\log \omega_i\}_{i=1}^n$, and $\mathbf{v} = \boldsymbol{\Omega}^{-a}\mathbf{x}/\|\boldsymbol{\Omega}^{-a}\mathbf{x}\|$ has unit norm.

Because M, and hence the derivative (1), vanishes when $a = 1$, our assertion is equivalent to the statement that $M > 0$ for $a \in [0,1)$. Define $f(\omega) = \omega^{2-2a}$ and set $f_i = f(\omega_i)$. Then

$$\begin{aligned} 2(1-a)M &= 2\sum_{i=1}^n (f_i \log f_i)v_i^2 - 2\sum_{i,j=1}^n (f_i \log f_j)v_i^2 v_j^2 \\ &= \sum_{i,j=1}^n (f_i - f_j)(\log f_i - \log f_j)v_i^2 v_j^2. \end{aligned} \tag{2}$$

Note that

$$(f_i - f_j)(\log f_i - \log f_j) > 0 \quad \text{whenever } \omega_i \neq \omega_j. \tag{3}$$

Furthermore, if any element x_i of $\mathbf{x}$ is 0, it can be removed (and the value of n reduced) without affecting the value of $\mathbf{x}'\boldsymbol{D}\boldsymbol{x} = \sum_{i=1}^n d_i x_i^2$ for a diagonal $\mathbf{D}$ and hence without affecting the value of $e_{a,\delta}$. Thus we can assume that all $x_i \neq 0$ and hence that all $v_i \neq 0$. This together with (3) ensures that (2) is positive (because $\omega_i \neq \omega_j$ at least once), thus establishing the monotonicity claim when $p = 1$.

Counterexample when $p > 1$. For general p the preceding method results in

$$\frac{d}{da} e_{a,\boldsymbol{\delta}} = -2\boldsymbol{\delta}'(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}\mathbf{M}(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}\boldsymbol{\delta},$$

where

$$\mathbf{M} = 2\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V} - (\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{V})(\mathbf{V}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V}) - (\mathbf{V}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V})(\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{V})$$

and $\mathbf{V} = \boldsymbol{\Omega}^{-a}\mathbf{X}(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}$ satisfies $\mathbf{V}'\mathbf{V} = \sum_{i=1}^n \mathbf{v}_i\mathbf{v}_i' = \mathbf{I}_p$. The analogue of (2) is that

$$2(1-a)\mathbf{M} = \sum_{i,j=1}^n (f_i - f_j)(\log f_i - \log f_j)\mathbf{v}_i\mathbf{v}_i'\mathbf{v}_j\mathbf{v}_j',$$

and the monotonicity claim is equivalent to the positive definiteness of $\mathbf{M}$. However, for $p > 1$, $\mathbf{M}$ often has at least one negative eigenvalue. In such cases the monotonicity properties will depend on the choice of $\boldsymbol{\delta}$. A particular counter-

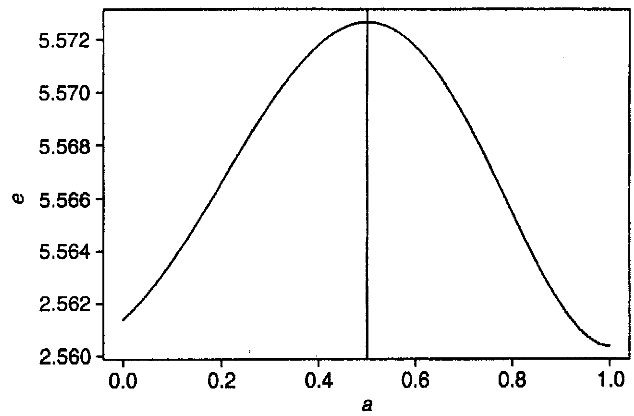

FIGURE 1. Plot of $e_{a,\boldsymbol{\delta}}$ against a with $\boldsymbol{\delta}$ chosen so that the maximum is attained at $a = 0.5$.

example to both suggestions of Professors Koenker and Portnoy has $n = 3$, $p = 2$, and

$$\boldsymbol{\Omega} = \begin{pmatrix} 5.4 & 0 & 0 \\ 0 & 2.9 & 0 \\ 0 & 0 & 1.8 \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} 0.2 & -1 \\ 1.2 & 0.4 \\ 0.3 & 0.5 \end{pmatrix}.$$

For these choices the eigenvalues of $\mathbf{M}$ are 1.49 and -0.003 when $a = 0.5$. We choose $\boldsymbol{\delta} = (1,0.26291)'$; then $e_{a,\boldsymbol{\delta}}$ is nonmonotone and is *maximized* at $a = 0.5$. See Figure 1.

This counterexample can be extended to all $n > 3$ by adding $n - 3$ rows of zeros to the bottom of $\mathbf{X}$ and continuing the diagonal of $\boldsymbol{\Omega}$ arbitrarily.

NOTE

1. An excellent solution was independently proposed by Geert Dhaene. A proof for the case $p = 1$ based on the convexity of Laplace transforms was also provided by Koenker and Portnoy, the posers of the problem.