On the stability of recursive least squares in the Gauss-Markov model

Salies, Evens

6 September 2004

On the stability of recursive least squares in the Gauss-markov model

Evens Salies

University of Paris 1 - Panthéon Sorbonne.
e.salies@caramail.com

Problem and Motivation

Consider the classical model $\boldsymbol{y}_n = \boldsymbol{X}_n\boldsymbol{\beta} + \boldsymbol{\varepsilon}_n$ where $\boldsymbol{X}_n$ is an $n \times p$ real matrix of fixed regressors, $\boldsymbol{y}_n$ $(n \times 1)$ a response vector, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown coefficients, $\mathrm{rk}(\boldsymbol{X}_n) = p$ for $n \geq p$. Let $\hat{\boldsymbol{\beta}}(n)$ denote the ordinary least squares estimate of $\boldsymbol{\beta}$ obtained from n observations, with $n \geq p$, and assume $\boldsymbol{\varepsilon}_n$ $(n \times 1)$ is a vector of non-observable random disturbances with expectation $\mathbf{0}$ and variance $\sigma^2 \boldsymbol{I}_n$.

An updating formula for $\hat{\boldsymbol{\beta}}(n+1)$ as a function of $\hat{\boldsymbol{\beta}}(n)$ is

$$\hat{\boldsymbol{\beta}}(n+1) - \boldsymbol{\beta} = \boldsymbol{W}^{-1}\boldsymbol{V}(\hat{\boldsymbol{\beta}}(n) - \boldsymbol{\beta}) + \boldsymbol{w},\ n = p, p+1, \ldots \tag{1}$$

where $\boldsymbol{V} \equiv \boldsymbol{X}_n'\boldsymbol{X}_n$, $\boldsymbol{W} \equiv \boldsymbol{X}_{n+1}'\boldsymbol{X}_{n+1}$, $\boldsymbol{w} \equiv \boldsymbol{W}^{-1}\boldsymbol{x}\varepsilon_{n+1}$, and $\boldsymbol{x}$ denotes the vector of new observations at the values of the explanatory variables. Eq. (1) arises for example in Kalman filtering and recursive least squares theories, where the unknown $\boldsymbol{\beta}$ is considered as time-varying states of dynamic system (see the discussion in Kianifard and Swallow, 1996) and $\boldsymbol{W}^{-1}\boldsymbol{V}$ is often developed as $\boldsymbol{I}_p - (1+c)^{-1}\boldsymbol{V}^{-1}\boldsymbol{x}\boldsymbol{x}'$; c equals $\boldsymbol{x}'\boldsymbol{V}^{-1}\boldsymbol{x}$.
This exercice provides some properties of $\boldsymbol{W}^{-1}\boldsymbol{V}$, with all its eigenvalues and eigenvectors. Let $\boldsymbol{A} \equiv \boldsymbol{W}^{-1}\boldsymbol{V}$ have eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_p$. Show that

(i) these eigenvalues are real, and that

(ii) $\lambda_1 = 1/(1+c),\ \lambda_2 = \lambda_3 = \cdots = \lambda_p = 1$.

Solution and Discussion

(i) $\boldsymbol{A}$ is the product between two real symmetric matrices. Let λ be an eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{u}+i\boldsymbol{v}$ an associated eigenvector, where $i^2=-1$. Then

$$\boldsymbol{A}(\boldsymbol{u}+i\boldsymbol{v})=\lambda(\boldsymbol{u}+i\boldsymbol{v}).$$

Premultiplying both sides of this equation with $\boldsymbol{W}$ leads to

$$\boldsymbol{V}(\boldsymbol{u}+i\boldsymbol{v})=\lambda\boldsymbol{W}(\boldsymbol{u}+i\boldsymbol{v}).$$

As $\boldsymbol{W}=\boldsymbol{V}+\boldsymbol{x}\boldsymbol{x}'$ therefore the previous equation becomes

$$(1-\lambda)\boldsymbol{V}(\boldsymbol{u}+i\boldsymbol{v})=\lambda\boldsymbol{x}\boldsymbol{x}'(\boldsymbol{u}+i\boldsymbol{v}).$$

Premultiply both sides with $(\boldsymbol{u}-i\boldsymbol{v})'$. Because of the symmetry of $\boldsymbol{V}$ we obtain

$$(1-\lambda)(\boldsymbol{u}'\boldsymbol{V}\boldsymbol{u}+\boldsymbol{v}'\boldsymbol{V}\boldsymbol{v})=\lambda((\boldsymbol{u}'\boldsymbol{x})^2+(\boldsymbol{v}'\boldsymbol{x})^2).$$

This implies that λ is real.

$\square$

(ii) The following determinant

$$\begin{aligned}|\boldsymbol{I}_p-\boldsymbol{A}| &= |\boldsymbol{I}_p-\boldsymbol{W}^{-1}\boldsymbol{V}|\\ &= |\boldsymbol{W}^{-1}(\boldsymbol{W}-\boldsymbol{V})|\\ &= |\boldsymbol{W}^{-1}|\cdot|\boldsymbol{x}\boldsymbol{x}'|\\ &= |\boldsymbol{W}|^{-1}\cdot 0\\ &= 0\end{aligned}$$

shows $\lambda=1$ is a root of the characteristic equation $|\lambda\boldsymbol{I}_p-\boldsymbol{A}|=0$. Now, let $\boldsymbol{z}$ be an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 1; therefore

$\boldsymbol{W}^{-1}\boldsymbol{V}\boldsymbol{z} = \boldsymbol{z}$ or $\boldsymbol{V}\boldsymbol{z} = \boldsymbol{W}\boldsymbol{z}$, which from the definition of $\boldsymbol{W}$ implies

$$\underset{(p\times 1)}{\mathbf{0}} = \boldsymbol{x}\boldsymbol{x}'\boldsymbol{z},$$

showing $\boldsymbol{z}$ is orthogonal to $\boldsymbol{x}$. Remaining eigenvalues of $\boldsymbol{A}$ are given using Wolkowicz and Styan's inequalities. We need trace($\boldsymbol{A}$).

$$\begin{aligned}
\operatorname{trace}(\boldsymbol{A}) &= \operatorname{trace}(\boldsymbol{W}^{-1}\boldsymbol{V}) \\
&= \operatorname{trace}(\boldsymbol{W}^{-1}(\boldsymbol{W} - \boldsymbol{x}\boldsymbol{x}')) \\
&= \operatorname{trace}(\boldsymbol{I}_p - \boldsymbol{W}^{-1}\boldsymbol{x}\boldsymbol{x}') \\
&= p - \boldsymbol{x}'\boldsymbol{W}^{-1}\boldsymbol{x}.
\end{aligned}$$

Moreover, premultiplying $\boldsymbol{W} = \boldsymbol{V} + \boldsymbol{x}\boldsymbol{x}'$ by $\boldsymbol{x}'\boldsymbol{W}^{-1}$ and postmultiplying it by $\boldsymbol{V}^{-1}\boldsymbol{x}$ implies $\boldsymbol{x}'\boldsymbol{W}^{-1}\boldsymbol{x} = c/(1+c)$. Consequently

$$\operatorname{trace}(\boldsymbol{A}) = p - c/(1+c),$$

and it can be shown $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ and $1/(1+c)$ the associated eigenvalue. Premultiplying $\boldsymbol{A}$ with $\boldsymbol{x}'$ gives

$$\begin{aligned}
\boldsymbol{x}'\boldsymbol{A} &= \boldsymbol{x}'(\boldsymbol{I}_p - \boldsymbol{W}^{-1}\boldsymbol{x}\boldsymbol{x}') \\
&= \boldsymbol{x}' - (\boldsymbol{x}'\boldsymbol{W}^{-1}\boldsymbol{x})\boldsymbol{x}' \\
&= (1 - \frac{c}{1+c})\boldsymbol{x}' \\
&= \frac{1}{1+c}\boldsymbol{x}'.
\end{aligned}$$

As $\boldsymbol{A}$ has real eigenvalues we can apply the inequalities of Wolkowicz and Styan reproduced in Magnus and Neudecker (1991, p. 239) to find the order of multiplicity of previously found eigenvalues:

$$\begin{aligned}
m - s(p-1)^{1/2} &\leq \lambda_1 \leq m - \frac{s}{(p-1)^{1/2}} \\
m + \frac{s}{(p-1)^{1/2}} &\leq \lambda_p \leq m + s(p-1)^{1/2},
\end{aligned}$$

where $m = (1/p)\text{trace}(\boldsymbol{A})$ and $s^2 = (1/p)\text{trace}(\boldsymbol{A}^2) - m^2$.

We obtain

$$1/(1+c) \quad \leq \lambda_1 \leq \quad 1 - \frac{2}{p}\frac{c}{1+c} \tag{2}$$

$$1 \quad \leq \lambda_p \leq \quad 1 + \frac{(p-2)}{p}\frac{c}{1+c}. \tag{3}$$

From Theorem 4 in Magnus and Neudecker (1991, p. 203),

$$\begin{array}{lcccl} & \lambda_1 & \leq \dfrac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \leq & \lambda_p \\ \Leftrightarrow & \lambda_1 & \leq \dfrac{\boldsymbol{x}'(\boldsymbol{I}_p - \boldsymbol{W}^{-1}\boldsymbol{x}\boldsymbol{x}')\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \leq & \lambda_p \\ \Leftrightarrow & \lambda_1 & \leq 1 - \boldsymbol{x}'\boldsymbol{W}^{-1}\boldsymbol{x} \leq & \lambda_p \\ \Leftrightarrow & \lambda_1 & \leq 1 - \dfrac{c}{1+c} \leq & \lambda_p \\ \Leftrightarrow & \lambda_1 & \leq \dfrac{1}{1+c} \leq & \lambda_p. \end{array}$$

Combination of Eq. (2) and this result gives $\lambda_1 = 1/(1+c)$, which implies equality holds on the left of Eq. (3), that is $\lambda_p = 1$ and the $p-1$ largest eigenvalues are equal (Magnus and Neudecker, 1991, p. 239).

$\square$

References

HARVEY, A. C. **(1990).** The Econometric analysis of time series. 2nd ed., Cambridge, MA: MIT Press.

KIANIFARD F., and SWALLOW, H. **(1996).** A review of the development and application of recursive residuals in linear models. *Journal of the American Statistical Association*, 91(433), pp. 391-400.

MAGNUS, J. R. and NEUDECKER, H. **(1991).** Matrix Differential Calculus with Applications in Statistics and Econometrics. John Wiley. Printed in Great Britain.