Journal of Economic Dynamics and Control 12 (1988) 43–48. North-Holland

THE FLEXIBLE LEAST SQUARES APPROACH TO TIME-VARYING LINEAR REGRESSION*

Robert KALABA and Leigh TESFATSION

University of Southern California, Los Angeles, CA 90089-0152, USA

1. Introduction

Consider an investigator who obtains noisy observations on a process and who wishes to learn about the actual sequence of states through which the process has passed. Suppose the investigator believes the process is adequately described by non-linear dynamic and measurement equations. However, he sees no way to justify the assignment of specific probabilistic properties to residual error terms.

In earlier studies [Kalaba and Tesfatsion (1980, 1981, 1985, 1986a)] a solution is developed for this state estimation problem assuming the non-linear dynamic and measurement equations are known up to a parameterization. Currently we are considering what might be done when the dynamic equations are unknown but the process state evolves only slowly over time. A smoothness prior is introduced in place of an explicit specification for the unknown dynamic equations governing the evolution of the process state.

In Kalaba and Tesfatsion (1986b) a special case of the latter problem is considered, namely, a variant of the well-known time-varying linear regression problem [Chow (1984)]. An investigator obtains noisy observations on a process which he believes can be adequately described by a linear regression model with a slowly evolving coefficient (state) vector. The actual dynamic equations governing the evolution of the coefficient vector are unknown and are proxied by a smoothness prior. Residual dynamic and measurement errors are anticipated to be small, but are otherwise unrestricted. The investigator wishes to estimate the sequence of time-varying coefficient vectors.

The 'flexible least squares (FLS) solution' proposed for this time-varying linear regression problem consists of all coefficient sequence estimates which attain the 'residual efficiency frontier' – i.e., which yield minimal pairs $(r_{\mathrm{D}}^2, r_{\mathrm{M}}^2)$ of squared residual dynamic error and measurement error sums, conditional on the given observations. A conceptually and computationally straightforward algorithm is developed which permits the exact sequential updating of

*A full version of this paper [Kalaba and Tesfatsion (1986b)] is available from the second author upon request.

the FLS estimates as the duration of the process increases and additional observations are obtained. Intrinsic relationships between the FLS and OLS solutions are established, and comparative robustness properties are determined for the case of a single unanticipated regime shift.

In Kalaba et al. (1987) a FORTRAN program is presented which implements the sequential updating algorithm for the FLS estimates. The program has been extensively tested and incorporates several validation checks that users can employ. All simulation experiments to date have been very encouraging. The qualitative time variation of the true coefficient vectors is effectively mimicked by the FLS estimates at each point along the residual efficiency frontier, despite noisy observations.

Section 2 describes the basic time-varying linear regression problem. Section 3 reviews the solution proposed for this problem in Kalaba and Tesfatsion (1986b). Simulation experiments are discussed in section 4.

2. The basic time-varying linear regression problem

Let N and K be arbitrary given integers satisfying $N \geq 2$ and $K \geq 1$. Suppose an investigator obtains noisy scalar observations $y_1, \ldots, y_N$ on a process over the time span $1, \ldots, N$. The investigator believes ex ante that the observations y_n have been generated in accordance with the linear regression model

$$y_n = x_n^T b_n + v_n, \qquad n = 1, \ldots, N, \tag{1}$$

where $x_n^T = (x_{n1}, \ldots, x_{nK})$ denotes a $1 \times K$ row vector of known exogenous regressors, $b_n = (b_{n1}, \ldots, b_{nK})^T$ denotes a $K \times 1$ column vector of unknown coefficients, and v_n denotes a scalar unobserved residual measurement error.

Suppose the investigator also believes ex ante that the residual measurement errors v_n are small and that the coefficient vectors b_n evolve slowly over time, if at all. The investigator's prior theoretical beliefs concerning the generation of the observations $y_1, \ldots, y_N$ might thus be cast in the following form:

Prior Measurement Specifications

$$y_n - x_n^T b_n \approx 0, \qquad n = 1, \ldots, N, \tag{2a}$$

Prior Dynamic Specifications

$$b_{n+1} - b_n \approx 0, \qquad n = 1, \ldots, N-1. \tag{2b}$$

A basic problem for the investigator is to determine whether there exists *any* coefficient sequence estimate $(b_1, \ldots, b_N)$ which satisfies the prior theoretical specifications (2a)–(2b) in an acceptable approximate sense for the realized observations $(y_1, \ldots, y_N)$. How might such a coefficient sequence estimate be found?

3. The flexible least squares approach

One possible approach is as follows. Associated with each possible coefficient sequence estimate $b = (b_1, \ldots, b_N)$ are two basic types of residual modelling error. First, b could fail to satisfy the prior measurement specifications (2a). Second, b could fail to satisfy the prior dynamic specifications (2b).

Suppose a cost is assigned to b for each type of error. For example, suppose the cost assigned to b for the first type of error is measured by the sum of squared residual measurement errors,

$$r_{\mathrm{M}}^2(b; N) = \sum_{n=1}^{N} \left[y_n - x_n^T b_n \right]^2, \tag{3}$$

and the cost assigned to b for the second type of error is measured by the sum of squared residual dynamic errors,

$$r_{\mathrm{D}}^2(b; N) = \sum_{n=1}^{N-1} [b_{n+1} - b_n]^T [b_{n+1} - b_n]. \tag{4}$$

Define the *(time N) residual possibility set* to be the collection $P(N)$ of all possible configurations $(r_{\mathrm{D}}^2(b; N), r_{\mathrm{M}}^2(b; N))$ of squared residual dynamic error and measurement error sums attainable at time N as b ranges over the space of possible coefficient sequence estimates. (See fig. 1.)

The greatest lower bound $P_{\mathrm{F}}(N)$ for the residual possibility set $P(N)$ gives the locus of minimal squared residual dynamic error and measurement error

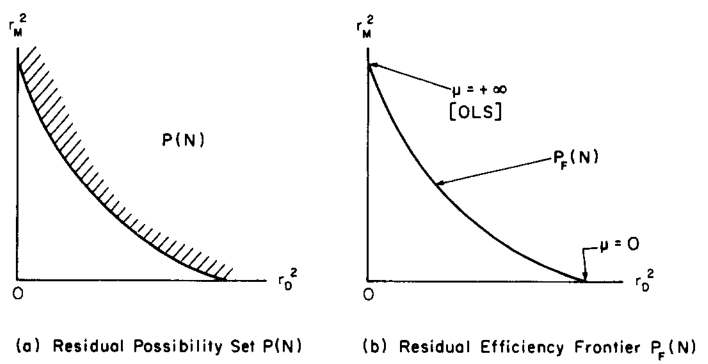

Fig. 1

sums attainable at time N, conditional on the given set of observations. Thus, given any coefficient sequence estimate yielding a point on $P_F(N)$, there exists no other coefficient sequence estimate which simultaneously lowers both types of error. Hereafter $P_F(N)$ will be referred to as the *(time N) residual efficiency frontier*.

How might the residual efficiency frontier be found? In analogy to the usual procedure for tracing out Pareto efficiency frontiers, a parameterized family of minimization problems is considered.

Thus, let $\mu \geq 0$ be given, and let each possible coefficient sequence estimate $b = (b_1, \ldots, b_N)$ be assigned an *incompatibility cost*,

$$C(b; \mu, N) = \mu r_D^2(b; N) + r_M^2(b; N), \tag{5}$$

consisting of the μ-weighted average of the associated measurement error and dynamic error sums (3) and (4). Expressing (5) in component form,

$$C(b; \mu, N) = \mu \left[\sum_{n=1}^{N-1} [b_{n+1} - b_n]^T [b_{n+1} - b_n] \right] + \sum_{n=1}^{N} [y_n - x_n^T b_n]^2. \tag{6}$$

As (6) indicates, the incompatibility cost function $C(b; \mu, N)$ generalizes the goodness-of-fit criterion function for ordinary least squares estimation by permitting the coefficient vectors b_n to vary over time.

Suppose the $K \times N$ regressor matrix $[x_1, \ldots, x_N]$ has rank K. Then, given any $\mu > 0$, there exists a unique coefficient sequence estimate $b = (b_1, \ldots, b_N)$ which minimizes the incompatibility cost function $C(b; \mu, N)$. [See Kalaba and Tesfatsion (1986b, sect. 4).] Let this unique minimizing estimate be denoted by

$$b^{FLS}(\mu, N) = \left(b_1^{FLS}(\mu, N), \ldots, b_N^{FLS}(\mu, N)\right). \tag{7}$$

If $\mu = 0$, let (7) denote any coefficient sequence estimate which minimizes the sum of squared residual dynamic errors $r_D^2(b; N)$ subject to $r_M^2(b; N) = 0$. Hereafter, (7) will be referred to as the *flexible least squares (FLS) solution, conditional on μ and N.*

Finally, let the sums of squared residual measurement errors and dynamic errors corresponding to the FLS solution (7) be denoted by

$$r_M^2(\mu, N) = r_M^2\left(b^{FLS}(\mu, N); N\right) \quad \text{and}$$

$$r_D^2(\mu, N) = r_D^2\left(b^{FLS}(\mu, N); N\right). \tag{8}$$

The residual efficiency frontier $P_F(N)$ then takes the parameterized form

$$P_F(N) = \{r_D^2(\mu, N), r_M^2(\mu, N) | 0 \leq \mu < \infty\}. \quad (9)$$

The residual efficiency frontier (9) is qualitatively depicted in fig. 1b. As μ approaches zero, the incompatibility cost function (5) ultimately places no weight on the prior dynamic specifications (2b) which require the coefficient vectors b_n to evolve only slowly. Thus, r_M^2 can generally be brought down close to zero, and the corresponding value for r_D^2 will be relatively large. As μ becomes arbitrarily large, the incompatibility cost function (5) places absolute priority on the prior dynamic specifications (2b); i.e., r_M^2 is minimized subject to $r_D^2 = 0$. The latter case coincides with ordinary least squares estimation in which a single $K \times 1$ coefficient vector is used to minimize the sum of squared residual measurement errors r_M^2.

4. Simulation experiments

In Kalaba et al. (1987) a FORTRAN program is presented which implements a conceptually and computationally straightforward algorithm for sequentially generating the FLS solution (7). As part of the program validation, various simulation experiments were performed involving linear, quadratic, and sinusoidal time variations in the true coefficient vectors $(b_1, \ldots, b_N)$.

In each experiment the residual efficiency frontier was adequately traced out by evaluating the residual error sums (8) over a rough grid of penalty weights μ increasing by powers of ten. No instability or other difficult numerical behavior was encountered. The qualitative time variation displayed by the true coefficient vectors was accurately reflected in the FLS estimates at each point along the residual efficiency frontier, despite noisy observations.

For example, an experiment was conducted with two-dimensional regressor vectors $[K = 2]$ and thirty observations $[N = 30]$ for which the components of the true time-n coefficient vector $b_n = (b_{n1}, b_{n2})$ were simulated to be sinusoidal functions of n. The first component, b_{n1}, moved through two complete periods of a sine wave over $(1, \ldots, N)$, and the second component, b_{n2}, moved through one complete period of a sine wave over $(1, \ldots, N)$. Each observation was perturbed by an additive normally distributed shock term representing about a five percent measurement error. When the penalty μ incurred for any time variation between successive coefficient estimates was set at $\mu = 1$, the FLS estimates b_{n1}^{FLS} and b_{n2}^{FLS} closely tracked the true coefficients b_{n1} and b_{n2}. As μ was increased from 1 to 100, the FLS estimates $(b_{n1}^{\text{FLS}}, b_{n2}^{\text{FLS}})$ at each time n were pulled inward toward the OLS solution $(0, 0)$; but the paths traced out by the FLS estimates over $(1, \ldots, N)$ still accurately reflected the two-period and one-period sinusoidal motions of the true coefficients b_{n1} and b_{n2}.

Recalling that 'smoothness' is the only prior information the FLS solution incorporates concerning the underlying patterns of change in the true regression coefficients, it is remarkable how well the FLS estimates have been able to reconstruct these patterns in experiment after experiment.

References

Chow, G., 1984, Random and changing coefficient models, Ch. 21 in: Z. Griliches and M. Intriligator, eds., Handbook of econometrics, Vol. 2 (Elsevier Science Publishers, Amsterdam) 1213–1245.

Kalaba, R. and L. Tesfatsion, 1980, A least-squares model specification test for a class of dynamic nonlinear economic models with systematically varying parameters, Journal of Optimization Theory and Applications 32, 538–567.

Kalaba, R. and L. Tesfatsion, 1981, An exact sequential solution procedure for a class of discrete-time nonlinear estimation problems, IEEE Transactions on Automatic Control AC-26, 1144–1149.

Kalaba, R. and L. Tesfatsion, 1985, Exact sequential filtering, smoothing, and prediction for nonlinear systems, Nonlinear Analysis: Theory, Methods, and Applications, forthcoming.

Kalaba, R. and L. Tesfatsion, 1986a, Sequential nonlinear estimation with nonaugmented priors, Journal of Optimization Theory and Applications, forthcoming.

Kalaba, R. and L. Tesfatsion, 1986b, Time-varying regression via flexible least squares, Modelling Research Group working paper no. 8633 (Department of Economics, University of Southern California, Los Angeles, CA).

Kalaba, R., N. Rasakhoo and L. Tesfatsion, 1987, A FORTRAN program for time-varying linear regression via flexible least squares, Modelling Research Group working paper no. 8730 (Department of Economics, University of Southern California, Los Angeles, CA).