Exact rational expectations, cointegration, and reduced rank regression

Søren Johansen,*
Department of Economics, University of Copenhagen
and
CREATES, University of Aarhus
and
Anders Rygh Swensen,†
University of Oslo and Statistics Norway

Abstract

We interpret the linear relations from exact rational expectations models as restrictions on the parameters of the statistical model called the cointegrated vector autoregressive model for non-stationary variables. We then show how reduced rank regression, Anderson (1951), plays an important role in the calculation of maximum likelihood estimation of the restricted parameters.

Keywords: Exact rational expectations, Cointegrated VAR model, Reduced rank regression
JEL Classification: C32

*Corresponding author: Søren Johansen, Department of Economics, University of Copenhagen, Studiestræde 6, DK-1455 Copenhagen K, Denmark. Email: sjo@math.ku.dk
The author acknowledges support from Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation.
†Department of Mathematics, University of Oslo, P.O. Box 1053, N-0316 Blindern, Oslo, Norway. Email: swensen@math.uio.no

1 Introduction

The purpose of this paper is to show how the technique of reduced rank regression, Anderson (1951), forms the basis for the calculation of Gaussian maximum likelihood estimation in the cointegrated vector autoregressive model, where the parameters are restricted by some exact rational expectations models.

Expectations play a major role in modern economics. Many variables such as long term interest rates are based on assumptions or expectations of future developments of other key economic variables. There exist many possibilities for modelling these expectations or forecasts. When the economic model is formulated by incorporating stochastic elements, a simple first choice is to let expectations mean probabilistic conditional expectations with respect to the information set of the model. This is what usually is called model based rational expectations, originally introduced by Muth (1961).

Rational expectations models specify relations between such conditional expectations of future values of some variables and past and present values of others. If no further stochastic terms are involved, *exact* rational expectations is the usual denomination. A well known example is the uncovered interest parity where the difference of the interest rates in two countries is supposed to equal the expected appreciation or depreciation of the exchange rate in the next period. Another example is a simple present value model for the price of stocks which is supposed to equal a discounted sum of future expected dividends.

To describe the simultaneous dynamic behavior of a moderate number of economic variables, a vector autoregressive model is a useful and often applied tool. In the context of exact rational expectations, an important aspect is that the conditional expectation of the variables *one step* ahead can easily be computed from the vector autoregressive model and expressed as linear combinations of present and past values. This means that exact rational expectations models imply restrictions on the coefficients of the vector autoregressive model. Thus the statistical model embeds the economic model of exact rational expectations. This makes it possible to test the validity of the economic model and also the specific values of coefficients which are assumed known by the rational expectations model. Test for linear relations based on the assumption of a stationary vector autoregressive model were developed by Hansen and Sargent (1981, 1991).

It is, however, a well recognized fact that many economic variables exhibit too large fluctuations for this behavior to be well captured by an assumption of stationarity. A reduced rank vector autoregressive model where one or several roots of the characteristic polynomial are equal to one, is a better alternative. This means that there are one or more stationary linear combinations of the variables. Using the fact that the coefficients of such cointegrating relations can be super-consistently estimated, Baillie (1989) generalized the approach of Hansen and Sargent (1981, 1991) by first using the estimated coefficients to transform the reduced rank vector autoregressive model to a model for stationary variables, and then test the restrictions in this model. This amounts to a two stage procedure: first transforming the variables to stationarity and then conducting inference in the remaining parameters in the model for stationary variables.

The method of reduced rank regression gives the possibility of obtaining explicit solutions to the nonlinear optimization problem posed by Gaussian maximum likelihood estimation in the cointegrated vector autoregressive model. In two previous papers, Johansen and Swensen (1999, 2004), we have shown this for some cointegrated vector autoregressive model models where the parameters are restricted by exact rational expectations models. We present here the solution for such a model where the rational expectations restrictions are allowed to be more general and where the cointegrating space in addition must satisfy some extra conditions. In this case a small modification of the argument is needed. Some examples are given below of this type of model with a homogeneity restriction imposed on the cointegrating space, so that the coefficients in the long-run relation add to zero.

One thus obtain a nesting of the models, where the reduced rank vector autoregressive model is the most general. The restrictions on the cointegration space represent the next level. Finally the last level consists of models satisfying the restrictions from the rational expectations hypothesis in addition to the restrictions on the long-run coefficients.

Likelihood ratio statistics of two common forms of additional restrictions can be calculated explicitly by reduced rank regression, see Johansen and Juselius (1990). One is where all the cointegrating relations satisfy the same linear constraints. The other is where some cointegration relations are assumed to be known. These are the restrictions considered here

There are a couple of points that should be stressed. The first is that only conditional expectations *one step* ahead are considered. Moreover, estimation is developed under the assumption that all coefficients describing the rational expectations relations are known up to some coefficient. This is clearly a limitation since they are often specified as containing some additional unknown parameters, such as discount factors in present value models, and they may enter nonlinearly. However, once we have an expression for the profile likelihood using reduced rank regression, a numerical optimization procedure can be used to find the maximum likelihood estimators of the remaining (few) parameters. Finally, we only consider vector autoregressive model models without constant and linear terms. This is done in order not to overload the exposition. However, combining the arguments in the present paper with the earlier papers Johansen and Swensen (1999, 2004) it should be fairly clear how such terms can be incorporated.

The paper is organized as follows. In the next section the model is introduced and the set of restrictions defined. The third section deals with the case that all cointegration vectors satisfy the same linear constraints in addition to the restrictions that follow from the rational expectations relations, and in the final section we treat the case that the additional restriction is that the cointegration space contains a known part.

We will use the usual notation that if a is an $n \times m,\ m < n$ matrix of full rank, then $\overline{a} = a(a'a)^{-1}$ and satisfies $a'\overline{a} = I_m$, and $a_\perp$ is an $n \times (n-m)$ matrix such that

$a'_\perp a = 0$ and the $n \times n$ matrix $(a, a_\perp)$ is nonsingular., finally $I_n = \bar{a}a' + \overline{a_\perp}a'_\perp$.

2 The restrictions implies by exact rational expectations

This section defines the cointegrated vector autoregressive model as the statistical model which is assumed to generate the data and formulates the parameter restrictions implied by the exact rational expectation hypothesis.

2.1 The cointegrated vector autoregressive model

Let the p-dimensional vectors of observations be generated according to the vector autoregressive model

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + \varepsilon_t,\ t = 1, \ldots, T \tag{1}$$

where $X_{-k+1}, \ldots, X_0$ are fixed and $\varepsilon_1, \ldots, \varepsilon_T$ are independent, identically distributed Gaussian vectors, with mean zero and covariance matrix Σ. We assume that $\{X_t\}_{t=1,2,\ldots}$ is $I(1)$ and that the $p \times r$ matrices α and β have full column rank r. This implies that X_t is non-stationary, ΔX_t is stationary, and that $\beta' X_t$ is stationary. It is the stationary relations between non-stationary processes and the interpretation as long-run relations, that has created the interest in this type of model in economics. We define the statistical model defined by (1).

$\mathcal{H}(r)$: The model is defined by equation (1) where α and β are $p \times r$ matrices and otherwise there are no further restrictions on the parameters. The number of identified parameters in the matrix $\alpha\beta'$ is $\#(\alpha\beta') = pr + r(p - r)$.

In the following we also assume that β is restricted either by homogeneity restrictions of the form $\beta = H\phi$, $(sp(\beta) \subset sp(H))$ or that some cointegrating vectors are known, $\beta = (b, \phi)$ $(sp(b) \subset sp(\beta))$. This defines two submodels of $\mathcal{H}(r)$.

$\mathcal{H}_1(r)$: The model is defined by equation (1) and the restriction $\beta = H\phi$, where H is a known $p \times s$ matrix of rank s, and ϕ is an $s \times r$ matrix of parameters, $r \leq s \leq p$. In this model $\#(\alpha\beta') = pr + r(s - r)$.

$\mathcal{H}_2(r)$: The model is defined by equation (1) and the restriction $\beta = (b, b_\perp\psi)$ where b is a known $p \times m$ matrix of rank m, and ψ is a $(p - m) \times (r - m)$ matrix of parameters, $m \leq r \leq p$. In this model $\#(\alpha\beta') = pr + (r - m)(p - r)$.

2.2 Estimation of the cointegrated vector autoregressive model, $\mathcal{H}(r)$, $\mathcal{H}_1(r)$, and $\mathcal{H}_2(r)$

It is well known, see Johansen (1996), that the Gaussian maximum likelihood estimator of β is calculated by reduced rank regression of ΔX_t on X_{t-1} corrected for the stationary regressors

$$\Delta X_{t-1}, \ldots, \Delta X_{t-k+1}.$$

Once β is determined, the other parameters are estimated by regression.

Model $\mathcal{H}_1(r)$ is estimated by reduced rank regression of ΔX_t on $H'X_{t-1}$ corrected for the stationary differences and their lags.

Finally model $\mathcal{H}_1(r)$ is estimated by noting that we already have the cointegrating relations b, and determine the $r-m$ remaining ones by reduced rank regression of ΔX_t on $b_\perp' X_t$ corrected for the stationary lagged differences and $b'X_{t-1}$.

2.3 The model for exact rational expectations and some examples

The model formulates a set of restrictions on the conditional expectation of X_{t+1} given the information $\mathcal{O}_t$ in the variables up to time t, which we write in the form

$\mathcal{RE}$: The model based exact rational expectations formulates relations for conditional expectations

$$E[c'\Delta X_{t+1}|\mathcal{O}_t] = \tau d' X_t + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t+1-i}. \tag{2}$$

Here $E[\cdot|\mathcal{O}_t]$ denotes the conditional expectation in the probabilistic sense of model (1), given the variables $X_1, \ldots, X_t$. The matrices $c(p \times q)$, $d(p \times n)$, $d_i(p \times n_i)$, $i = 1, \ldots, \ell$ are *known* full rank matrices, and $\tau(q \times n)$, $\tau_i(q \times n_i), i = 1, \ldots, \ell$ are parameters. We assume that $n \leq q$ and $\ell \leq k$.

We give next three examples of rational expectations models, where the last two examples have the property that the implied cointegration relation satisfies a homogeneity restriction, which correspond to the models we analyse in this paper.

Example 1 Let $X_t = (Y_t, y_t)'$ consist of the price of stock, Y_t, at the end of period t and of the dividends, y_t, paid during the period t. The present value model entails, see Campbell and Shiller (1987), that the price of stock can be expressed as a discounted sum of the expected future values of dividends given the information at time t, i.e.

$$Y_t = \sum_{i=1}^{\infty} \delta^i E[y_{t+i}|\mathcal{O}_t].$$

Equivalently, after some rearrangement, we find the equation

$$E[\Delta(Y_{t+1} + y_{t+1})|\mathcal{O}_t] = (\delta^{-1} - 1)Y_t - y_t. \tag{3}$$

For known δ, this has the form (2) with

$$c' = (1, 1),\ d' = (\delta^{-1} - 1, -1),\ \ell = 0,$$

and the coefficient τ is known and equal to one. Obviously (3) implies that if (Y_t, y_t) are $I(1)$ variables, so that $(\Delta Y_t, \Delta y_t)$ is stationary, then $(\delta^{-1} - 1)Y_t - y_t$ is a cointegrating relation. Moreover no further lags are needed in order to describe the conditional expectation. In this example the coefficient δ is not known. For given value of δ, we can apply regression to concentrate out all other parameters and we are left with a function of one parameter to be optimized. ■

Example 2 We let $X_t = (e_{12t}, i_{1t}, i_{2t})$ where e_{12t} is the exchange rate between two countries and i_{1t} and i_{2t} are the long term bond interest rates. The uncovered interest parity assumes that

$$E[\Delta e_{12t+1}|\mathcal{O}_t] = i_{1t} - i_{2t}.$$

This has the form (2) with $c' = (1, 0, 0)$, $d' = (0, 1, -1)$, $\ell = 0$, and the parameter τ is known and equal to one. In this case an $I(1)$ model for X_t would imply that $(0, 1, -1)$ is a cointegration vector. Because in theory the conditional expectation only depends on i_{1t} and i_{2t}, this implication can be investigated by testing restrictions on the short-run dynamics. ■

Example 3 This example is taken from Boug et al. (2006). Let pa_t denote log export price of machinery and equipment, pf_t the corresponding import price, and let mc_t denote the marginal cost, all in logarithms. The linear quadratic adjustment cost model minimizes the cost incurred by not hitting the target, pa_t^*, on the one hand and the cost resulting from changing prices on the other hand:

$$E_t[\sum_{j=0}^{\infty} \beta^j (\theta(pa_{t+j} - pa_{t+j}^*)^2 + (\Delta pa_{t+j})^2)].$$

This implies the first order condition

$$\Delta pa_t = \beta E_t[\Delta pa_{t+1}] - \theta(pa_t - pa_t^*).$$

The target is assumed to satisfy

$$pa_t^* = \gamma pf_t + (1 - \gamma)mc_t,$$

so that the theoretical model has the implication that

$$\Delta pa_t = \beta E_t[\Delta pa_{t+1}] - \theta(pa_t - \gamma pf_t - (1 - \gamma)mc_t).$$

For $I(1)$ variables this implies that $(1, -\gamma, -1+\gamma)$ is a cointegration vector, which evidently satisfied the homogeneity restriction. If we express the relation as

$$E_t[\Delta pa_{t+1}] = \frac{\theta}{\beta}(pa_t - \gamma pf_t - (1-\gamma)mc_t) + \frac{1}{\beta}\Delta pa_t \tag{4}$$

we have an example of (2) with

$$c' = (1,0,0),\ d' = (1, -\gamma, -1+\gamma),\ \ell = 1,\ d_1' = (1,0,0).$$

In this example the coefficients $\tau = \theta/\beta$ and $\tau_1 = 1/\beta$ have to be estimated.

2.4 Combining the exact rational expectations and the vector autoregressive models

We now combine the exact rational expectations and the vector autoregressive models, $H_1(r)$ and $H_2(r)$, and express the exact rational expectations model (2) as restrictions on the coefficients of the statistical model (1). Taking the conditional expectation of $c'\Delta X_{t+1}$ given $X_1, \ldots, X_t$, we get using (1),

$$c'E[\Delta X_{t+1}|\mathcal{O}_t] = c'\alpha\beta' X_t + \sum_{i=1}^{k} c'\Gamma_i \Delta X_{t+1-i}.$$

Equating this expression to (2) implies that the following conditions must be satisfied

$$c'\alpha\beta' = \tau d',\ c'\Gamma_i = \tau_i d_i',\ i = 1, \ldots, \ell,\ c'\Gamma_i = 0,\ i = \ell+1, \ldots, k.$$

This can be summarized as

Proposition 1 *The exact rational expectations restrictions (2) give the following restrictions on the parameters of model (1):*

$$c'\alpha\beta' = \tau d', \tag{5}$$

$$c'\Gamma_i = \tau_i d_i',\ i = 1, \ldots, \ell,\ c'\Gamma_i = 0,\ i = \ell+1, \ldots, k. \tag{6}$$

Note that (5) implies that $\tau d'\beta_\perp = 0$, so that when $n \leq q$ we find $\bar{\tau}'\tau d'\beta_\perp = d'\beta_\perp = 0$, and hence $sp(d) \subset sp(\beta)$ and therefore $n \leq r$.

We define two submodels of $H_1(r)$ and $H_2(r)$ respectively which satisfy the restrictions in RE, where the parameters $\tau, \{\tau_i\}_{i=1}^{\ell}$ are unrestricted and have to be estimated. We mention in section 3.2.1 how the estimation procedure should be modified if one or more of the coefficients are known.

$\mathcal{H}_1^\dagger(r)$: The model is a submodel of $\mathcal{H}_1(r)$ which satisfies the restrictions (5) and (6).

$\mathcal{H}_2^\dagger(r)$: The model is a submodel of $\mathcal{H}_2(r)$ which satisfies the restrictions (5) and (6).

Assumption (5) implies that $sp(d) \subset sp(\beta)$, so that in $\mathcal{H}_1^\dagger(r)$ it holds that $sp(d) \subset sp(\beta) \subset sp(H)$, whereas in $\mathcal{H}_2^\dagger(r)$ we have $sp(b, d) \subset sp(\beta)$.

When estimating models $\mathcal{H}_1^\dagger(r)$ and $\mathcal{H}_2^\dagger(r)$ it is convenient to use a parametrization of freely varying parameters. Such a parametrization is given for the two models in sections 3 and 4 together with an analysis of the estimation problem, which is solved by regression and reduced rank regression.

3 The same restrictions on all β

We first give a representation in terms of freely varying parameters of the matrix $\alpha\beta'$, when restricted by $\beta = H\phi$ and $c'\alpha\beta' = \tau d'$, see (5). Next we show how estimation of $\mathcal{H}_1^\dagger(r)$ can be performed by regression and reduced rank regression.

3.1 A reparametrization of $\mathcal{H}_1^\dagger(r)$

Proposition 2 *Let d and H satisfy $d = Hd^*$ for some d^* and that $q \geq n$ and $p - q \geq r - n$. Then the restrictions*

$$\beta = H\phi \tag{7}$$

and

$$c'\alpha\beta' = \tau d' \tag{8}$$

hold if and only if

$$\alpha\beta' = \bar{c}\tau d' + c_\perp\theta d' + c_\perp\kappa\varsigma' d_\perp^{*\prime} H', \tag{9}$$

where κ is $(p-q)\times(r-n)$ and ς' is $(r-n)\times(s-n)$. This implies that the number of identified parameters in $\alpha\beta'$ is

$$\#(\alpha\beta') = qn + (p-q)r + (r-n)(s-r).$$

Proof. Assume first that (9) holds. Multiplying (9) by c' we find that $c'\alpha\beta = \tau d'$ which proves (8), and by multiplying (9) by $\bar{\alpha}'$ and $H_\perp$ we find from $d'H_\perp = 0$, that $\beta' H_\perp = 0$, so that $\beta = H\phi$ which proves (7). Moreover the row space of $\alpha\beta'$ is spanned by $\tau d' = \tau d^{*\prime} H'$ and $\varsigma' d_\perp^{*\prime} H'$, which are linearly independent and of rank n and $r - n$ respectively so that β has rank r. This proves (7) and (8).

Next assume that (7) and (8) are true. Note that by multiplying by $\bar{H}$ we find that (8) implies that $c'\alpha\phi' = \tau d^{*\prime}$. We construct $d_\perp^*$, $s \times (s-n)$, and find

$$(c, \overline{c_\perp})'\alpha\phi'(\bar{d}^*, \overline{d_\perp^*}) = \begin{pmatrix} \tau & 0 \\ \overline{c_\perp}'\alpha\phi'\bar{d}^* & \overline{c_\perp}'\alpha\phi'\overline{d_\perp^*} \end{pmatrix}.$$

We let $\theta = \overline{c_\perp}'\alpha\phi'\bar{d}^*$. We define the $r \times n$ matrix $\xi = \alpha' c \bar{\tau}$, and $\xi_\perp$, $r \times (r-n)$. Then, because $\xi'\phi' d_\perp^* = \bar{\tau}' c'\alpha\phi' d_\perp^* = \bar{\tau}'\tau d^{*\prime} d_\perp^* = d^{*\prime} d_\perp^* = 0$, we find

$$\overline{c_\perp}'\alpha\phi'\overline{d_\perp^*} = \overline{c_\perp}'\alpha(\bar{\xi}\xi' + \overline{\xi_\perp}\xi_\perp')\phi'\overline{d_\perp^*} = (\overline{c_\perp}'\alpha\overline{\xi_\perp})(\xi_\perp'\phi'\overline{d_\perp^*}) = \kappa\varsigma',$$

where $\kappa = \overline{c_\perp}'\alpha\overline{\xi_\perp}$ is $(p-q)\times(r-n)$, and $\varsigma' = \xi_\perp'\phi'\overline{d_\perp^*}$ is $(r-n)\times(s-n)$. Hence, using $(\bar{d}^*, \overline{d_\perp^*})^{-1} = (d^*, d_\perp^*)'$ and $(c, \overline{c_\perp}')^{-1} = (\bar{c}, c_\perp)'$, we find

$$\alpha\phi' = \bar{c}\tau d^{*\prime} + c_\perp\theta d^{*\prime} + c_\perp\kappa\varsigma' d_\perp^{*\prime}.$$

When multiplied by H' we have proved (9). ∎

3.2 Estimation of model $\mathcal{H}_1^{\dagger}(r)$

We analyse the Gaussian likelihood by the equation for $c'\Delta X_t$ and the conditional equation for $c_{\perp}'\Delta X_t$ given $c'\Delta X_t$.

3.2.1 The equation for $c'\Delta X_t$

We find from (1) and (6) the equation for $c'\Delta X_t$

$$c'\Delta X_t = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\varepsilon_t. \tag{10}$$

If the parameters τ and $\{\tau_i\}_{i=1}^{\ell}$, are unknown, they and $\Sigma_{cc} = c'\Sigma c$ are estimated by regression of $c'\Delta X_t$ on $d'X_{t-1}, \{d_i'\Delta X_{t-i}\}_{i=1}^{\ell}$. If τ_ℓ, say, is known and equal to τ_ℓ^0, then (10) is estimated by regression of $c'\Delta X_t - \tau_\ell^0 d_\ell' \Delta X_{t-\ell}$ on $d'X_{t-1}, \{d_i'\Delta X_{t-i}\}_{i=1}^{\ell-1}$. In the following we state the results for the model where the parameters $\tau, \tau_1, \ldots, \tau_\ell$ are unknown. The number of parameters in the regression coefficients is $q(n + n_1 + \ldots + n_\ell)$.

3.2.2 The equation for $\overline{c_\perp}'\Delta X_t$ conditional on $c'\Delta X_t$

This equation follows from (1) and (9)

$$\begin{aligned} \overline{c_\perp}'\Delta X_t &= \rho c'\Delta X_t + (\theta - \rho\tau) d' X_{t-1} + \kappa \varsigma' d_\perp^{*\prime} H' X_{t-1} \\ &\quad + \sum_{i=1}^{\ell} (\overline{c_\perp}'\Gamma_i - \rho \tau_i d_i') \Delta X_{t-i} + \sum_{i=\ell+1}^{k} \overline{c_\perp}'\Gamma_i \Delta X_{t-i} + (\overline{c_\perp}' - \rho c')\varepsilon_t, \end{aligned}$$

where $\rho = \overline{c_\perp}'\Sigma c (c'\Sigma c)^{-1} = \Sigma_{c_\perp c}\Sigma_{cc}^{-1}$ and the variance of the residuals is

$$\Sigma_{c_\perp c_\perp . c} = \overline{c_\perp}'\Sigma \overline{c_\perp} - \overline{c_\perp}'\Sigma c (c'\Sigma c)^{-1} c'\Sigma \overline{c_\perp}.$$

The parameters $(\tau,\ \{\tau_i\}_{i=1}^{\ell},\ \Sigma_{cc})$ and

$$(\rho, \theta - \rho\tau, \kappa, \varsigma, \{\overline{c_\perp}'\Gamma_i - \rho\tau_i d_i'\}_{i=1}^{\ell}, \{\overline{c_\perp}'\Gamma_i\}_{i=\ell+1}^{k}, \Sigma_{c_\perp c_\perp . c})$$

are freely varying. This implies that estimation of the conditional equation can be performed by reduced rank regression of $\overline{c_\perp}'\Delta X_t$ on $d_\perp^{*\prime} H' X_{t-1}$ corrected for the stationary regressors

$$(c'\Delta X_t, d'X_{t-1}, \{\Delta X_{t-i}\}_{i=1}^{k}).$$

This determines the remaining $r - n$ cointegrating relations. We summarize the result in the next proposition.

Proposition 3 *The maximum likelihood estimators for models $\mathcal{H}_1^{\dagger}(r)$ and $\mathcal{H}_1(r)$ can be calculated by regression and reduced rank regression, and the likelihood ratio statistic $-2logLR(\mathcal{H}_1^{\dagger}(r)|\mathcal{H}_1(r))$ is asymptotically distributed as $\chi^2(f)$ with degrees of freedom $f = qs + kpq - q(n + n_1 + \ldots + n_\ell)$.*

Proof. The estimation result is given above and the asymptotic results follow from the general results about inference in the cointegrated vector autoregressive model, see Johansen (1996). The number of unknown parameters in the conditional mean in the model $\mathcal{H}_1(r)$ is $pr + r(s-r) + kp^2$. In model $\mathcal{H}_1^{\dagger}(r)$ the number of parameters is $(p-q)r + (r-n)(s-r) + q(n+n_1+\ldots+n_\ell) + k(p-q)p$ which gives the degrees of freedom. ■

4 Some β assumed known

We first give a representation in turns of freely varying parameters of the matrix $\alpha\beta'$ when restricted by $\beta = (b, b_\perp\phi)$ and $c'\alpha\beta' = \tau d'$. Next we show that estimation of $\mathcal{H}_2^{\dagger}(r)$ can be conducted by regression and reduced rank regression.

4.1 A reparametrization of $\mathcal{H}_2^{\dagger}(r)$

We start by investigating the relation between the spaces spanned by d and b. The spaces may or may not overlap and the matrix $b'd_\perp$ could be of any rank from zero to $\min(m,(p-n))$, depending on the relative positions of the spaces. We decompose $b'\overline{d_\perp} = uv'$, where u, $m \times n'$, and v, $(p-n)\times n'$, and define the orthogonal decomposition of $\mathbb{R}^p$

$$(d, b_1, b_2) = (d, d_\perp v, \overline{d_\perp} v_\perp)$$

of dimensions $(n, n', p-n-n')$ respectively. Note that

$$b = d\bar{d}'b + d_\perp \overline{d_\perp}' b = d\bar{d}'b + b_1 u',$$

and that $sp(b,d) = sp(d, b_1)$, and finally that b_2 span the orthogonal complement of $sp(d,b)$.

Proposition 4 *Let b and d be given and assume $r \geq rank(b,d)$, and that $q \geq n$ and $p-q \geq r-n$. Then the restrictions*

$$\beta = (b, \psi) \tag{11}$$

and

$$c'\alpha\beta' = \tau d' \tag{12}$$

hold if and only if

$$\alpha\beta' = \bar{c}\tau d' + c_\perp\theta_1 d' + c_\perp\theta_2 b_1' + c_\perp\kappa\varsigma' b_2', \tag{13}$$

where $b'\overline{d_\perp} = uv', b_1 = d_\perp v$, and $b_2 = \overline{d_\perp} v_\perp$, and κ is $(p-q)\times(r-n-n')$ and ς' is $(r-n-n')\times(p-n-n')$. This implies that the number of identified parameters is

$$\#(\alpha\beta') = qn + (p-q)r + (r-n-n')(p-r).$$

Proof. Assume first that (13) holds, then

$$c'\alpha\beta' = c'\bar{c}\tau d' + c'(c_\perp\theta_1 d' + c_\perp\theta_2 b_1' + c_\perp\kappa\varsigma' b_2') = \tau d'$$

which shows (12). Next we prove that b is part of the cointegrating space, see (11). We have from (12) that $\tau d'\beta_\perp = 0$, and because $q \geq n$ we find that $d'\beta_\perp = 0$, and hence $d \in sp(\beta)$. We want to show that $b_1 \in sp(\beta)$, which together with $d \in sp(\beta)$, shows that also $b \in sp(\beta)$.

We find from (12) that because $d'\beta_\perp = 0$, it holds that

$$0 = \overline{c_\perp}'\alpha\beta'\beta_\perp = \theta_2 b_1'\beta_\perp + \kappa\varsigma' b_2'\beta_\perp.$$

Now κ is $(p-q)\times(r-n-n')$, and because $p-q \geq r-n \geq r-n-n'$ we find that $\kappa_\perp$ is $(p-q)\times(p-q-r+n+n')$, so that

$$\kappa_\perp'\theta_2 b_1'\beta_\perp = 0.$$

The matrix $\kappa_\perp'\theta_2$ is $(p-q-r+n+n')\times n'$, and because $p-q-r+n+n' \geq n'$ we therefore find $b_1'\beta_\perp = 0$. This shows that $b_1 \in sp(\beta)$ and that $r \geq n+n' = rank(b,d)$ which proves (11).

Finally we want to show that (11) and (12) implies (13). We investigate the matrix

$$(c, \overline{c_\perp})'\alpha\beta'(\bar{d}, \overline{b_1}, \overline{b_2}) = \begin{pmatrix} \tau & 0 & 0 \\ \overline{c_\perp}'\alpha\beta'\bar{d} & \overline{c_\perp}'\alpha\beta'\overline{b_1} & \overline{c_\perp}'\alpha\beta'\overline{b_2} \end{pmatrix}.$$

We define $\theta_1 = \overline{c_\perp}'\alpha\beta'\bar{d}$ and $\theta_2 = \overline{c_\perp}'\alpha\beta'\overline{b_1}$. We know that $sp(d,b) = sp(d,b_1) \subset sp(\beta)$, and can write $(d,b_1) = \beta\xi$ for some ξ. Then, because $\xi'\beta'\overline{b_2} = (d,b_1)'\overline{b_2} = 0$ we have

$$\overline{c_\perp}'\alpha\beta'\overline{b_2} = \overline{c_\perp}'\alpha(\bar{\xi}\xi' + \overline{\xi_\perp}\xi_\perp')\beta'\overline{b_2} = (\overline{c_\perp}'\alpha\overline{\xi_\perp})(\xi_\perp'\beta'\overline{b_2}),$$

and we define $\kappa = \overline{c_\perp}'\alpha\overline{\xi_\perp}$, $(p-q)\times(r-n+n')$, and $\zeta' = \xi_\perp'\beta'\overline{b_2}$, $(r-n-n')\times(p-n-n')$. This implies that

$$\alpha\beta' = \bar{c}\tau d' + c_\perp\theta_1 d' + c_\perp\theta_2 b_1' + c_\perp\kappa\zeta' b_2'.$$

■

4.2 Estimation of model $\mathcal{H}_2^\dagger(r)$

We analyse the Gaussian likelihood by the equation for $c'\Delta X_t$ and the conditional equation for $c_\perp'\Delta X_t$ given $c'\Delta X_t$.

4.2.1 The equation for $c'\Delta X_t$

The marginal equation for $c'\Delta X_t$ is given by (10) and the parameters $(\tau, \{\tau_i\}_{i=1}^{\ell}, c'\Sigma c)$ are estimated by regression.

4.2.2 The equation for $c_\perp' \Delta X_t$ conditional on $c'\Delta X_t$

The conditional equation for $c_\perp' \Delta X_t$ given $c'\Delta X_t$ is found from (1) and (13)

$$\overline{c_\perp}' \Delta X_t = \rho c' \Delta X_t + (\theta_1 - \rho\tau) d' X_{t-1} + \theta_2 b_1' X_{t-1} + \kappa \varsigma' b_2' X_{t-1}$$
$$+ \sum_{i=1}^{\ell} (\overline{c_\perp}' \Gamma_i - \rho \tau_i d_i') \Delta X_{t-i} + \sum_{i=\ell+1}^{k} \overline{c_\perp}' \Gamma_i \Delta X_{t-i} + (\overline{c_\perp}' - \rho c') \varepsilon_t,$$

where $\rho = \Sigma_{c_\perp c} \Sigma_{cc}^{-1}$. The parameters $(\tau,\ \{\tau_i\}_{i=1}^{\ell},\ \Sigma_{cc})$ and

$$(\rho, \theta_1 - \rho\tau, \theta_2, \kappa, \varsigma, \{\overline{c_\perp}' \Gamma_i - \rho \tau_i d_i'\}_{i=1}^{\ell}, \{\overline{c_\perp}' \Gamma_i\}_{i=\ell+1}^{k}, \Sigma_{c_\perp c_\perp . c})$$

are variation independent. This shows the conditional model can be again analysed by reduced rank regression of $\overline{c_\perp}' \Delta X_t$ on $b_2' X_{t-1}$ corrected for the stationary regressors

$$(c' \Delta X_t, d' X_{t-1}, b_1' X_{t-1}, \{\Delta X_{t-i}\}_{i=1}^{k}).$$

Thus we exploit that in $\mathcal{H}_2^{\dagger}(r)$ there are $n + n'$ known cointegrating vectors spanning $sp(d, b)$ and the remaining $r - n - n'$ are determined by reduced rank regression.

Proposition 5 *The maximum likelihood estimators for both models $\mathcal{H}_2^{\dagger}(r)$ and $\mathcal{H}_2(r)$ can be calculated by regression and reduced rank regression, and the likelihood ratio statistic $-2logLR(\mathcal{H}_2^{\dagger}(r)|\mathcal{H}_2(r))$ is asymptotically distributed as χ^2 with degrees of freedom equal to $rq + (n + n' - m)(p - r) + kpq - q(n + n_1 + \ldots + n_\ell)$.*

Proof. The estimation result is given above and the asymptotic results follow from the general results about inference in the cointegrated vector autoregressive model, see Johansen (1996). The number of unknown parameters in the model $\mathcal{H}_2^{\dagger}(r)$ which are contained in (13) is $(p - q)r + (r - n - n')(p - r) + qn$. In addition there are $k(p - q)p$ free parameters in the matrices $\overline{c_\perp}' \Gamma_i, i = 1, \ldots, k$ and $q(n_1 + \ldots + n_\ell)$ free parameters in $\tau_i, i = 1, \ldots, \ell$. The number of parameters in model $\mathcal{H}_2(r)$ is $pr + (r - m)(p - r) + kp^2$, and the degrees of freedom of the test is the difference. ∎

5 An application

We consider the model discussed in Example 3. We take quarterly seasonally unadjusted data from the Norwegian economy over the period 1978:1 to 1998:4. The data is analysed in Boug et al. and it was found that a VAR with four lags, seasonal dummies and an unrestricted constant, gave a satisfactory description of the data, if a dummy which is +1 in 1979:1 and -1 in 1979:2 is included in the equation for Δpa_t. The model had constant parameters over the period considered. The cointegrating rank was found to be one, and the cointegration vector estimated in $\mathcal{H}(1)$ is

$$pa = \hat{\alpha}_0 + \underset{(0.227)}{0.525} pf + \underset{(0.230)}{0.585} mc, \ \log L_{\max}(\mathcal{H}(1)) = 931.48$$

with the standard error in parenthesis.

Model $\mathcal{H}_1(1)$ with the homogeneity restriction gives estimates

$$pa = \hat{\alpha}_0 + \underset{(0.344)}{0.645} pf + (1 - 0.645)mc, \ \log L_{\max}(\mathcal{H}_1(1)) = 931.24,$$

and the test of the homogeneity restriction, $-2 \log LR(\mathcal{H}_1(1)|\mathcal{H}(1)) = 0.49$, is asymptotically $\chi^2(1)$ with a $p-$value of 0.49. Figure 1 shows the likelihood profile as a function of γ for model $\mathcal{H}_1(1)$.

Next we want to test $\mathcal{H}_1^{\dagger}(1)$, which has in addition the zero restrictions on the short-run dynamics: $e_1'\Gamma_1 e_2 = e_1'\Gamma_1 e_3 = 0$, and $e_1'\Gamma_2 = e_1'\Gamma_3 = 0$, a total of eight restrictions. The equations for $\mathcal{H}_1^{\dagger}(r)$ become

$$\begin{aligned} \Delta pa_t &= \alpha_1 d' x_{t-1} + e_1'\Gamma_1 e_1 \Delta pa_{t-1} + e_1'\Phi_1 D_t + \varepsilon_{1t}, \\ \Delta pf_t &= \alpha_2 d' x_{t-1} + e_2'\Gamma_1 \Delta x_{t-1} + e_2'\Gamma_2 \Delta x_{t-2} + e_2'\Gamma_3 \Delta x_{t-3} + e_2'\Phi D_t + \varepsilon_{2t}, \\ \Delta mc_t &= \alpha_3 d' x_{t-1} + e_3'\Gamma_1 \Delta x_{t-1} + e_3'\Gamma_2 \Delta x_{t-2} + e_3'\Gamma_3 \Delta x_{t-3} + e_3'\Phi D_t + \varepsilon_{3t}. \end{aligned} \tag{14}$$

Here we have introduced D_t for the deterministic terms and note that $\alpha_1 = \theta/\beta$ and $e_1'\Gamma_1 e_1 = 1/\beta$ vary freely. The vector $d' = (1, -\gamma, -1 + \gamma)$ depends on an unknown parameter so we can apply the results in section 3 to optimize for fixed γ, and then afterwards optimize over γ. The analysis of the equations (14) for fixed γ can in this case be conducted by regression because there is only one cointegrating relation:

$$\begin{aligned} &1 : \text{Regress } \Delta pa_t \text{ on } d'x_{t-1}, \Delta pa_{t-1}, D_t, \\ &2 : \text{Regress } \Delta pf_t \text{ on } \Delta pa_t, d'x_{t-1}, \Delta x_{t-1}, \Delta x_{t-2}, \Delta x_{t-3}, D_t, \\ &2 : \text{Regress } \Delta mc_t \text{ on } \Delta pf_t, \Delta pa_t, d'x_{t-1}, \Delta x_{t-1}, \Delta x_{t-2}, \Delta x_{t-3}, D_t. \end{aligned}$$

From the estimated variances $(\hat{\sigma}_{11}^2(\gamma), \hat{\sigma}_{22.1}^2(\gamma), \hat{\sigma}_{33.1,2}^2(\gamma))$ we find the likelihood profile

$$\log L_{\max}(\mathcal{H}_1^{\dagger}(r))(\gamma) = -\frac{1}{2} T \log(\hat{\sigma}_{11}^2(\gamma)\hat{\sigma}_{22.1}^2(\gamma)\hat{\sigma}_{33.1,2}^2(\gamma)).$$

The maximum likelihood estimate becomes

$$pa = \hat{\alpha}_0 + \underset{(0.504)}{0.534} pf + (1 - 0.534)mc, \ \log L_{\max}(\mathcal{H}_1^{\dagger}(r)) = 914.75,$$

and the standard error is found by a quadratic approximation to the likelihood profile. Figure 1 also shows the likelihood profile as a function of γ for model (14) with the restrictions on the short-run dynamics, and in both cases with the homogeneity restriction imposed on d.

A comparison of the maximized log likelihood values shows that

$$-2 \log LR(\mathcal{H}_1^{\dagger}(r)|\mathcal{H}_1(r)) = 2(931.24 - 914.75) = 32.98$$

which is a test of eight restrictions on the coefficients of lagged differences in the equation for Δpa_t. The test statistic is asymptotically distributed as $\chi^2(8)$ and is thus highly significant, which shows that although the number of cointegrating relations and the homogeneity restriction assumed by the theory are in accordance with the data, the simple dynamic structure assumed by the economic model is not.

Thus the example shows that, in this case, the rational expectation formulation of the economic theory is too simple to give a good description of the data.

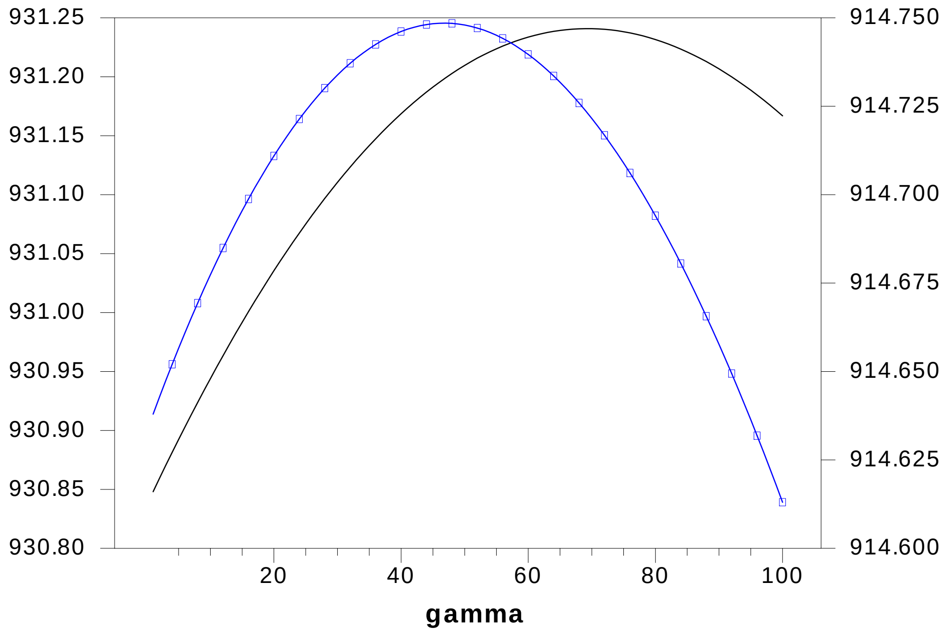

Figure 1: The profile likelihoods for model $\mathcal{H}_1(1)$ (line and left hand scale) gives $\hat{\gamma} = 0.645$ and $\log L_{\max}(\mathcal{H}_1(1)) = 931.24$. For model $\mathcal{H}_1^{\dagger}(r)$ (squares and right hand scale) we find $\hat{\gamma} = 0.534$ and $\log L_{\max}(\mathcal{H}_1^{\dagger}(r)) = 914.75$.

6 Conclusion

Reduced rank regression was developed more than 50 years ago and has been used in limited information maximum likelihood estimation ever since. In cointegration analysis it found a new domain of applicability. We have shown here that some exact rational expectations model, when embedded in the cointegrated vector autoregressive model, can be estimated by a combination of regression and reduced rank regression. Not all such models have a simple solution, but even in more complicated models, reduced rank regression is a simple way of eliminating many parameters in the vector autoregressive model.

References

[1] Anderson, T.W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics* **22**, 327–351.

[2] Baillie, R.T. (1989). Econometric tests of rationality and market efficiency. *Econometric Reviews* **8**, 151–186.

[3] Boug, P., Cappelen, Å. and A.R. Swensen (2006). Expectations and regime robustness in price formation: evidence from vector autoregressive models and recursive methods. *Empirical Economics* **31,** 821–845.

[4] Campbell, J.Y. and R.J. Shiller (1987). Cointegration tests and present value models. *Journal of Political Economy* **95**, 1062–1088.

[5] Hansen, L.P. and T.J. Sargent (1981). Exact linear rational expectations models: Specification and estimation. Research Department Staff Report 71, Federal Reserve Bank of Minneapolis.

[6] Hansen, L.P. and T.J. Sargent (1991). Exact linear rational expectations models: Specification and estimation. In *Rational Expectations Econometrics.* Hansen, L.P. and T.J. Sargent eds., Westview Press, Boulder.

[7] Johansen, S. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford.

[8] Johansen, S. and K. Juselius (1990). Maximum Likelihood Estimation and Inference on Cointegration — with Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics* **52**, 169–210.

[9] Johansen, S. and A.R. Swensen (1999). Testing exact rational expectations in cointegrated vector autoregressive models. *Journal of Econometrics* **93**, 73–91.

[10] Johansen, S. and A.R. Swensen (2004). More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term. *Econometrics Journal* **7**, 389–397.

[11] Muth, J.F. (1961). Rational expectations and the theory of price movements. *Econometrica* **29**, 315–335.