

Common features in vector autoregressive models¹

Caratteristiche comuni in modelli autoregressivi vettoriali

Paolo Paruolo
Department of Economics
University of Insubria, Varese²

Riassunto: Questo lavoro presenta una breve rassegna su caratteristiche comuni in modelli autoregressivi vettoriali, con particolare riferimento al caso di trend e cicli comuni. Si mostra come la presenza di queste caratteristiche comuni corrisponde a restrizioni di rango di particolari trasformazioni lineari dei coefficienti autoregressivi. Tali restrizioni possono essere testate mediante tecniche di regressione di rango ridotto introdotte in Anderson (1951), basate sulle correlazioni canoniche. Tali tecniche danno soluzione esplicita al problema di massima verosimiglianza gaussiana. Le distribuzioni asintotiche risultano di tipo diverso per differenti caratteristiche. Si riportano inoltre alcune estensioni al caso di integrazione stagionale e a restrizioni di tipo dinamico. Una possibile strategia di modellizzazione è illustrata con un'applicazione su dati finanziari USA.

Keywords: Common features, Cointegration, Common trends, Common cycles, Vector autoregressive processes, $I(1)$, Reduced rank regression, canonical correlations.

1. Introduction

The notion of common factors is a classical idea in statistics; during the last two decades it has received new momentum in the study of time series thanks to the introduction of cointegration, see Engle and Granger (1987). The importance of this notion has been recently recognized by the Nobel Price Award 2003 in economics to Engle and Granger.

The notion of cointegration is a special case of common features, as defined by Engle and Kozicki (1993) and Vahid and Engle (1993). This definition also covers common stationary cycles as a special case. In Section 2 we review these two types of common features, cointegration and common stationary cycles, for vector autoregressive processes (VAR). We emphasize that both types of cofeature result in reduced rank restrictions on some linear combinations of the autoregressive coefficient matrix. This connection is presented in Section 3.

We also review some extensions of this concept to dynamic linear combinations in Section 4. These extensions include polynomial cofeatures, see Cubadda and Hecq (2001), seasonal cointegration, see Cubadda (1999), Cubadda (2001b), Cubadda (2001a), nested reduced rank regression, see Ahn and Reinsel (1988), scalar component models, see Tiao and Tsay (1989), co-integrated processes of order 2, see Paruolo (2002). We do not treat the extensions in full detail but refer to the above references and Paruolo (2003).

The Gaussian likelihood analysis of reduced rank regression is briefly reviewed in Section 5. The asymptotic analysis in the case of cointegration and common cycles is very different. In the cointegration case, the cointegration parameters are consistent at a

¹Via Ravasi 2, 21100 Varese. Email: paolo.paruolo@uninsubria.it

²Partial financial support from MIUR Cofin 2002 and FAR grants is gratefully acknowledged.

rate T , where T is the sample size, i.e. they are superconsistent. The common cycle case involves only stationary processes, and leads to usual χ^2 asymptotics at the rate $T^{1/2}$.

We also discuss the interplay of these two notions within the same system. This allows to characterize systems with common trends and cycles. The system may be transformed using the estimated cointegration parameters in order to analyze common cycles. Fixing the cointegration parameters at their maximum likelihood (ML) estimates allows to make asymptotic inference on common cycles in these transformed models as if the cointegration parameters are known, thanks to the superconsistency of the cointegration parameters, see Johansen (1996). This property is not shared by the transformation proposed in Gonzalo and Granger (1995). In Section 6 an illustration on financial US time series is presented. Section 7 concludes and gives direction of future development.

2. Definitions

In this section we review the notion of common trends and cycles, following Johansen (1996) Chapter 3 for cointegration and Engle and Kozicki (1993), Vahid and Engle (1993) for common cycles. Consider a generic vector time series $W_t := (W_{1t} : \dots : W_{mt})'$, where W_{it} is univariate, which is a function of a vector X_t generated by a VAR(k), $k \geq 1$, system of the type

$$X_t = \sum_{i=1}^k A_i X_{t-i} + \epsilon_t, \quad (1)$$

where X_t and ϵ_t are $p \times 1$. L and $\Delta := 1 - L$ are the lag and difference operators, where negative powers of Δ indicate summation.³ W_t is taken to be e.g. equal to ΔX_t or to linear combinations of X_t and X_{t-1} . ϵ_t is assumed to be a martingale difference sequence with respect to $\mathcal{I}_t$, the sigma-field generated by X_{t-i} , $i \geq 0$, $E(\epsilon_t | \mathcal{I}_{t-1}) = 0$, with $E(\epsilon_t \epsilon_t' | \mathcal{I}_{t-1}) = \Omega$ p.d.. We take this as a definition of ‘innovation process’ and of ‘unpredictability’, i.e. we call any process W_t with $E(W_t | \mathcal{I}_{t-1}) \neq 0$ ‘predictable’. Stationary predictable processes are called ‘cycles’.

A process W_t is called an integrated process of order d at frequency zero, $I(d)$, when $\Delta^d W_t$ is stationary and $\Delta^{d-1} W_t$ is non-stationary.⁴ Let $\mathcal{F}$ be the class of univariate processes with a specific characteristic F, called feature, to be defined shortly. We define F a ‘feature’ if the following holds:

1. if $W_{it} \in \mathcal{F}$ (to be read ‘ W_{it} has feature F’) then $aW_{it} \in \mathcal{F}$ for $a \in \mathbb{R} \setminus \{0\}$;
2. if $W_{it} \notin \mathcal{F}$, (to be read ‘ W_{it} does not have feature F’) $i = 1, 2$ then $a_1 W_{1t} + a_2 W_{2t} \notin \mathcal{F}$ for $a_i \in \mathbb{R}$, $i = 1, 2$;
3. if $W_{1t} \notin \mathcal{F}$ and $W_{2t} \in \mathcal{F}$ then $W_{1t} + W_{2t} \in \mathcal{F}$.

Note that a feature F is a dominant property, in the sense that any linear combinations of one process within $\mathcal{F}$ and another one not in $\mathcal{F}$ gives a process within $\mathcal{F}$. A vector time series is called ‘with feature F’ if any of its components is in $\mathcal{F}$. It is easy to verify that $I(1)$ -ness and predictability are possible features in the above sense. $F=I$ indicates the former case. A ‘cycle’ is defined as any predictable stationary process. Also cyclicity is a feature in the above sense which is indicated with $F=C$.

³Deterministic terms are also included in the VAR equations in Paruolo (2003).

⁴For a definition of integration at seasonal frequencies see Cubadda (1999), Cubadda (2001a), Cubadda (2001b).

We next say that a features F is common if W_t has feature F and there exist a $b_i \in \mathbb{R}^m \setminus \{0\}$ such that $b_i' W_t$ does not have feature F . b_i is called a ‘cofeature vector’ and the set $\mathcal{B}_F := \{b : b \in \mathbb{R}^m, b' W_t \notin \mathcal{F}\}$ is called ‘cofeature space’. It is simple to verify that $\mathcal{B}$ is a linear vector subspace of $\mathbb{R}^m$, because it is closed under scalar multiplication and vector addition by 2. above. A basis b_F of $\mathcal{B}_F$ is called a ‘cofeature matrix’, and the dimension of the space $\mathcal{B}_F$ is called ‘cofeature dimension’ or ‘cofeature rank’.

When $F=I$ one obtains the case of co-integration, while when $F=C$ one obtains the case of co-cyclicity, cc for short. $\mathcal{B}_I$ is called the cointegration space, $\mathcal{B}_C$ the cc space.

3. Cofeature restrictions

We first consider the case of cointegration. It is convenient to rewrite $A(L)$ as $A(L) = -\Pi L + \Gamma(L)\Delta$ by expanding the VAR polynomial around $z = 1$; this gives the alternative equivalent representation

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \epsilon_t = \Pi X_{t-1} + \gamma V_t + \epsilon_t \quad (2)$$

where $\gamma := (\Gamma_1 : \dots : \Gamma_{k-2})$, $V_t := (\Delta X'_{t-1} : \dots : \Delta X'_{t-k+1})'$; let also $\Gamma := I - \sum_{i=1}^{k-1} \Gamma_i = \Gamma(1)$. Johansen (1988) proved Granger’s representation theorem for VAR processes, under the assumption that the only possible source of non-stationarity comes from the root $z = 1$ of the characteristic polynomial $\det(A(z))$. Under this assumption, necessary and sufficient conditions for X_t to be I(1) are given by

$$\Pi = \alpha \beta' \quad (3)$$

where α and β are $p \times r$ matrices of rank r , and $\alpha'_\perp \Gamma \beta_\perp$ is of full rank $p - r$. Here $\alpha_\perp$ represents a basis of the orthogonal complement of the column space spanned by α , $\text{col}(\alpha)$, and similarly for $\beta_\perp$. We will call these conditions the ‘I(1) conditions’.

Note that cointegration implies the reduced rank restriction (3) on a matrix coefficient in the linear model (2). Under (3) the representation (2) becomes

$$\Delta X_t = (\alpha : \gamma)(X'_{t-1} \beta : V'_t)' + \epsilon_t = \Psi U_t + \epsilon_t, \quad (4)$$

which is called the Equilibrium Correction Mechanism, EqCM, see Hendry (1995).

Under the I(1) conditions, it can be proved, see Johansen (1988) or Paruolo (2003), that $Y_t := (Y'_{0t} : Y'_{1t})' := (X'_t \beta : \Delta X'_t \beta_\perp)'$ follows a stationary VAR(k) process

$$Y_t = \sum_{i=1}^k A_i^\circ Y_{t-i} + \epsilon_t^\circ = \Psi^\circ U_t + \epsilon_t^\circ \quad (5)$$

where the coefficient of $Y_{1,t-k}$ is restricted to be zero, $\epsilon_t^\circ := B' \epsilon_t$ and $B := (\beta : \beta_\perp)$ is a nonsingular transformation of the innovations ϵ_t . (5) is called the Equilibrium Dynamics form, EqD. The vector of variables U_t can be taken to contain $(Y'_{0,t-1} : V'_t)'$ or equivalently $(Y'_{t-1} : \dots : Y'_{t-k+1} : Y'_{0,t-k})'$, which is the set of stationary regressors in (2); for a proof see Paruolo (2003).

The stationarity of (5) implies that $\beta' X_t$ are cointegration relations and that $\beta = b_I$ is a basis of $\mathcal{B}_I$. Inversion of the VAR (5) gives the MA(∞) representation $Y_t = C^\circ(L) \epsilon_t^\circ =$

$C^\circ(L)B'\epsilon_t$, where $C^\circ(1)$ is of full rank p . Since the moving average impact matrix $C^\circ(1)$ plays an important role in the analysis of common cycles, we call Y_t is an I(0) process of rank p .

From this representation one can deduce a MA(∞) representation for ΔX_t , simply selecting appropriate combinations of the $C_i^\circ B'$ matrices. This gives

$$\Delta X_t =: C^*(L)\epsilon_t = C\epsilon_t + C_0(L)\Delta\epsilon_t, \quad (6)$$

where $C := \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp$ is of rank $p - r$. This shows that ΔX_t is an I(0) process with rank $p - r$. Cumulating the MA representation for ΔX_t one finds the so-called common trends representation for X_t

$$X_t = C \sum_{i=1}^t \epsilon_i + C_0(L)\epsilon_t + m_0, \quad (7)$$

where the I(1) common $p - r$ trends are contained in $C \sum_{i=1}^s \epsilon_i$ and m_0 depends on initial values.

One can note that the rank of the MA impact matrix, i.e. the I(0) rank in the terminology above, plays an upper bound on the dimension ℓ of the co-cyclicity space $\mathcal{B}_C$. In fact in order for b_C to be a cc-cofeature matrix for a generic vector MA process $W_t = C_W(L)\epsilon_t$, one needs $b'_C C_{W,i} = 0$, $i = 1, 2, \dots$. Decomposing $C_W(L)$ as $C_W(L) = C_W(1) + C_{W0}(L)\Delta$,⁵ this in turn implies that $C_{W0,i} = 0$ for all i and $b'_C C_{W,i}(1) = q'$, where q is required to be of full column rank, otherwise the innovation process $q'\epsilon_t$ would fail to have p.d. second moments $q'\Omega q$. Hence $\ell \leq \text{rank}(q) \leq \min(\text{rank}(b_C), \text{rank}(C_{W,i}(1)))$. Since b_C can be chosen freely, this means that $\ell \leq \text{rank}(C_{W,i}(1))$. This restriction is more stringent for the choice $W_t = \Delta X_t$ with I(0) rank $p - r$ than for the choice $W_t = Y_t$ with I(0) rank p . For a more detailed characterization of the b_C in the former case see Paruolo (2003).

Engle and Kozicki (1993) and Vahid and Engle (1993) have proposed to look for co-cyclicity relations in ΔX_t ; in Paruolo (2003) I have proposed to also look for co-cyclicity in Y_t . The two options are complementary. Both type of cc-cofeatures are characterized by the reduced rank condition. Let $\Psi^{(\cdot)}$ indicate Ψ in (4) for the choice $W_t = \Delta X_t$, and $\Psi^{(\cdot)} = \Psi^\circ$ in (5) for the choice $W_t = Y_t$. In order for $b'_C W_t$ not to depend on $b'_C \Psi^{(\cdot)} U_t$, one needs $b'_C \Psi^{(\cdot)} = 0$, i.e. that $\text{rank}(\Psi^{(\cdot)}) = p - \ell$, where ℓ is the cc dimension. The reduced rank condition $\text{rank}(\Psi^{(\cdot)}) = p - \ell$ can be written $\Psi^{(\cdot)} = \varphi \tau'$, with φ and τ of full column rank $s := p - \ell$. In this case the cofeature matrix b_C can be chosen equal to $\varphi_\perp$. This is parallel to the rank decomposition (3) of Π . One thus sees that different cc-cofeature matrices correspond to bases of the left null space of $\Psi^{(\cdot)}$, while the cointegration matrix corresponds to a basis of the row space of the matrix Π .

4. Extensions

There are many possible extensions of the idea of cofeature matrix as defined in Section 1. In this section we review one possible extension, which corresponds to an increase of the set of variables W_t , allowed to contain variables selected from lags of X_t . As an

⁵This is always possible because all MA processes $C_W(L)\epsilon_t$ considered here have geometrically decreasing coefficients, and hence $C_W(z)$ is analytic for $|z| < 1 + \delta$, $\delta > 0$.

example of this extension, assume for instance that $W_t = (W'_{1t} : W'_{2t})'$ where W_{2t} is measurable with respect to F_{t-1} and W_t is measurable with respect to F_t . For instance one can have $W_{1t} = Y_t$, $W_{2t} = Y_{t-1}$. Partition also $b_C = (b'_{1C} : b'_{2C})'$ conformably with W_t . The cofeature relation $b'_C W_t = \sum_{i=1}^2 b'_{iC} W_{it}$ is equal $\sum_{i=1}^2 b'_{iC} Y_{t-i} = a(L)Y_t$, $a(L) := b'_{1C} + b'_{2C}L$, i.e. it represents a polynomial relation in the lag operator L .

It is simple to see that this extension does not add anything to the idea of cointegration for I(1) processes. Take in fact $W_{1t} = X_t$ and $W_{2t} = X_{t-1}$. Since the dimension of the cointegration space is invariant with respect to linear bijective transformations and $X_t = \Delta X_t + X_{t-1}$, this choice is equivalent to choosing $W_{1t} = \Delta X_t$ and $W_{2t} = X_{t-1}$, where ΔX_t is I(0). Hence nothing changes when increasing the set of variables to include more variables than X_{t-1} , because the cointegration relation are the same ones as for X_{t-1} plus the ones that select the stationary variables ΔX_t .⁶

Restrict hence attention to extensions of the cc-cofeature. A polynomial cc-cofeature linear combination of ΔX_t , $a(L)\Delta X_t$, has been termed ‘polynomial serial correlation common features’ by Cubadda and Hecq (2001). Similar extensions can also be applied to Y_t in place of ΔX_t . We call both these cases ‘unpredictable polynomial linear combinations’, UP.

UP restrictions for Y_t or ΔX_t correspond to other reduced ranks restrictions. In fact define $\epsilon_t^* := C_{W,0}\epsilon_t$, where $C_{W,0} = I$ for $W_t = \Delta X_t$ and $C_{W,0} = B'$ for $W_t = Y_t$, with covariance matrix of ϵ_t^* indicated as $\Omega^* := C_{W,0}\Omega C'_{W,0}$. Let $Z_{2t} := W_{2t}$, $Z_{0t} := W_{1t}$, where Z_{0t} satisfies

$$Z_{0t} = \varsigma Z_{1t} + \Phi Z_{2t} + \epsilon_t^*, \quad (8)$$

and Z_{1t} , Z_{2t} depend on lagged ϵ_t^* s. Then a necessary and sufficient condition for b_C to be a cofeature matrix for $(W'_{1t} : W'_{2t})' = (Z'_{0t} : Z'_{2t})'$ is that ς is of reduced rank, $\varsigma = \varphi\tau'$, with φ and τ of full column rank. In this case the cofeature matrix has representation

$$b'_C := (b'_{1C} : b'_{2C}) = (\varphi'_\perp : \varphi'_\perp \Phi) \quad \text{and} \\ b'(W'_{1t} : W'_{2t})' = \varphi'_\perp W_{1t} + \varphi'_\perp \Phi W_{2t} = \varphi'_\perp \epsilon_t^*,$$

see Paruolo (2003) for a proof. We note that all cases in Section 2 are special cases of the restrictions of this type. Other special cases are the ‘weak form’ of common features proposed in Hecq *et al.* (2002) for I(1) systems; the idea is that some elements in ΔX_t inherit the cyclic part included in deviations from equilibrium in $Y_{0,t-1}$. Other special cases are nested reduced rank regression of Ahn and Reinsel (1988) and scalar components of Tiao and Tsay (1989).

The UP extension allows also to ascertain which choice of W_{2t} would make the cc-cofeature properties of W_t equivalent⁷ for the choices $W_{1t} = Y_t$ and $W_{1t} = \Delta X_t$. It can be proved that cofeature properties of $U_{1t} := (\Delta X'_t : Y'_{0,t-1})'$ and $U_{2t} := (Y'_t : Y'_{0,t-1})'$ are equivalent, see again Paruolo (2003).

5. Likelihood analysis

In order to define the log-likelihood function $\mathcal{L}$, assume that ϵ_t is i.i.d. $N(0, \Omega)$; $2\mathcal{L}$ is then proportional to $-T(\det(\Omega^*) + \text{tr}(\Omega^{*-1}M_{\epsilon^*\epsilon^*}))$ where $M_{\epsilon^*\epsilon^*} := T^{-1} \sum_{t=1}^T \epsilon_t^* \epsilon_t^{*'}$ and ϵ_t^* is

⁶On the contrary, this extension makes sense in the case of I(2) processes, and it gives rise to the so-called ‘multicointegration’ or ‘polynomial cointegration’ case, see Haldrup (1998).

⁷In the following we say that the cofeature properties of two sets of variables are equivalent if they have the same cofeature rank and the cofeature matrices are linearly related.

defined in $Z_{0t} = \varsigma Z_{1t} + \Phi Z_{2t} + \epsilon_t^*$. The cofeature restriction both for cc and cointegration are of the type

$$H(s) : \quad \varsigma = \varphi \tau' \quad (9)$$

within (8), for appropriate choice of variables. For instance for the cointegration case $\varphi = \alpha$, $\tau = \beta$ and $Z_{0t} = \Delta X_t$, $Z_{1t} = X_{t-1}$, $Z_{2t} = V_t$. We observe that in the case of cointegration τ is the parameter of interest, a basis of $\mathcal{B}_I$, while in the case of cc, the parameter of interest is $\varphi_\perp$, a basis of $\mathcal{B}_C$. The $H(s)$ model is analyzed by the reduced rank regression $\text{RRR}(Z_{0t}, Z_{1t} | Z_{2t}, 1)$. The likelihood function $\mathcal{L}$ is maximized by considering the following eigenvalue problem

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0 \quad (10)$$

with eigenvalues $\lambda_1 > \dots \lambda_i > \dots > \lambda_p$ and associated eigenvectors v_i , where $S_{ij} := M_{ij} - M_{I1} M_{22}^{-1} M_{2j}$, $M_{ij} := T^{-1} \sum_{t=1}^T Z_{it} Z_{jt}'$, $i, j = 0, 1, 2$. The LR test statistic for hypothesis (9) of $H(s)$ versus $H(p)$ about the rank of ς is given by

$$Q_1(s) := -T \sum_{i=s+1}^p \ln(1 - \lambda_i).$$

This test, called the ‘trace test’, is asymptotically $\chi^2((j-s)(p-s))$ for the case of cc under the null while it presents a squared multivariate Dickey Fuller distribution for the cointegration case, see Johansen (1996). The limit distributions in the cc case are unaffected if β in the Y_t transformation is known or it is estimated, as long as r is correctly selected, see Paruolo (2003), unlike the transformation proposed in Gonzalo and Granger (1995) which also involves α .

In order to estimate the dimension r of the cointegration space and the complement $s = p - \ell$ to the dimension of the cc-cofeature space, the sequence of $Q_1(s)$ test can be set up from $s = 0$ to $s = p$. Johansen (1992) proved that this sequence selected the correct dimension with limit probability equal to 1 minus the significance level of each test in the sequence; moreover the probability of under-selecting r or s vanishes asymptotically. His proof goes through also in the case of cc-cofeatures. This procedure has thus probability equal to the nominal size of each test to over-select the cointegration rank, and of under-selecting the dimension of the cc-cofeature space.

Eq. (10) provides also the ML parameter estimates for given dimension s . In particular $\hat{\tau} = (v_1 : \dots : v_s)$ and

$$\begin{aligned} \hat{\varphi} &= S_{01} \hat{\tau} (\hat{\tau}' S_{11} \hat{\tau})^{-1}, & \hat{\varsigma} &= \hat{\varphi} \hat{\tau}' = S_{01} \hat{\tau} (\hat{\tau}' S_{11} \hat{\tau})^{-1} \hat{\tau}', \\ \hat{\Phi} &= (M_{02} - \hat{\varsigma} M_{12}) M_{22}^{-1}, & \hat{\Omega}^* &= S_{00} - S_{01} \hat{\tau} (\hat{\tau}' S_{11} \hat{\tau})^{-1} \hat{\tau}' S_{10}, \end{aligned} \quad (11)$$

where $\hat{\tau}$ is normalized by $\hat{\tau}' S_{11} \hat{\tau} = I_s$, $\hat{\tau}' S_{10} S_{00}^{-1} S_{01} \hat{\tau} = \text{diag}(\lambda_1, \dots, \lambda_s) =: \Lambda_1$.

In order to identify parameters, it is convenient to normalize $\hat{\tau}$ by the just-identifying restrictions $\hat{\tau}_c := \hat{\tau} (c' \hat{\tau})^{-1}$, where c is a known matrix of the same dimensions of τ , such that $c' \tau$ is a square nonsingular matrix, see (Johansen, 1996, Section 5.2). The choice of $\hat{\varphi}$ obtained by substituting $\hat{\tau}_c$ in place of $\hat{\tau}$ in (11) is given by ${}_c \hat{\varphi} := \hat{\varsigma} c$, which satisfies ${}_c \hat{\varphi} \hat{\tau}_c' = \hat{\varsigma}$. In the following we use the just-identifying normalization $\hat{\varphi}_{\perp a_\perp} := \hat{\varphi}_\perp (a'_\perp \hat{\varphi}_\perp)^{-1}$ also for the estimator of $\varphi_\perp$.

Consider a cc-hypothesis of the form $K' \text{vec}(\bar{a}' \varphi_{\perp a_\perp}) = j$, where K has rank h . Note that $\bar{a}' \varphi_{\perp a_\perp}$ are the un-normalized coefficients in $\varphi_\perp$ when it is identified via $\varphi_{\perp a_\perp}$. The

hypothesis can be tested via the Wald statistic

$$J_1 := T(K' \text{vec}(\bar{a}' \hat{\varphi}_{\perp a \perp}) - j)' \left(\left(\hat{\varphi}_{\perp a \perp}' \hat{\Omega}^* \hat{\varphi}_{\perp a \perp} \right)^{-1} \otimes (a' \hat{\zeta} S_{11} \hat{\zeta}' a) \right) (K' \text{vec}(\bar{a}' \hat{\varphi}_{\perp a \perp}) - j).$$

The test J_1 has an asymptotic $\chi^2(h)$ limit distribution. A similar Wald test can be defined for the corresponding hypothesis $K' \text{vec}(c_{\perp}' \tau_c) = j$ on the cointegration parameters; this is also asymptotically χ^2 distributed.

A special case of the previous hypothesis is given by

$$H_0 : \quad \varphi_{\perp} = H\phi,$$

which accommodate exclusion restrictions for all columns of $\varphi_{\perp}$ simultaneously. Here H is $p \times h$, $h \geq \ell$. Under this restriction the likelihood function is maximized by solving

$$|\lambda^* H' S_{00} H - H' S_{01} S_{11}^{-1} S_{10} H| = 0,$$

with LR test statistic $Q_2 := T \left(\sum_{i=s+1}^p \ln(1 - \lambda_i) - \sum_{i=h-p+s+1}^h \ln(1 - \lambda_i^*) \right)$. Under the null this test is asymptotically $\chi^2((p-h)(p-s))$. A similar hypothesis on τ can also be solved with a dual eigenvalue problem interchanging 0 and 1; this also gives a χ^2 test also in the cointegration case.

Consider now the case where b_C is (partly) known. Let K be a known $p \times h$ matrix of rank $h \leq \ell$, and consider the hypothesis that K is a submatrix of b_C , $b_C = (K, b_2^*)$, i.e.

$$H_0 : \quad K' \zeta = 0. \quad (12)$$

A Wald test of (12) can be based on the unrestricted maximum likelihood estimates $\hat{\zeta} := S_{01} S_{11}^{-1}$, $\hat{\Omega}^* := S_{00.1} := S_{00} - S_{01} S_{11}^{-1} S_{10}$, and equals

$$J_2 := T \text{tr} \left((K' S_{00.1} K)^{-1} K' \hat{\zeta} S_{11} \hat{\zeta}' K \right) = T \text{tr} \left((K' S_{00.1} K)^{-1} K' S_{01} S_{11}^{-1} S_{10} K \right). \quad (13)$$

This test is asymptotically $\chi^2(hj)$. The corresponding LR test of (12) in $H(s)$, labelled Q_3 , is found by solving the eigenvalue problem

$$|\lambda^{\circ} K'_{\perp} S_{00.K} K_{\perp} - K'_{\perp} S_{01.K} S_{11.K}^{-1} S_{10.K} K_{\perp}| = 0, \quad (14)$$

with eigenvalues $\lambda_1^{\circ} > \dots > \lambda_{p-h}^{\circ}$. The test of (12) in $H(s)$ is given by

$$Q_3 := T \left(\sum_{i=1}^s \ln(1 - \lambda_i^{\circ}) - \sum_{i=1}^s \ln(1 - \lambda_i) \right) \quad (15)$$

where $S_{00.1} := S_{00} - S_{01} S_{11}^{-1} S_{10}$. The Q_3 test is asymptotically $\chi^2(sh)$. Analogous result hold for $K' \zeta' = 0$ substituting φ with τ , and subscript 0 with 1, see Johansen (1996).

Other tests are also possible, see Johansen (1996), Paruolo (2003). All LR test obviously also provide restricted estimates as by-products. This allows to perform a specification search, as illustrated in the application summarized in the next section.

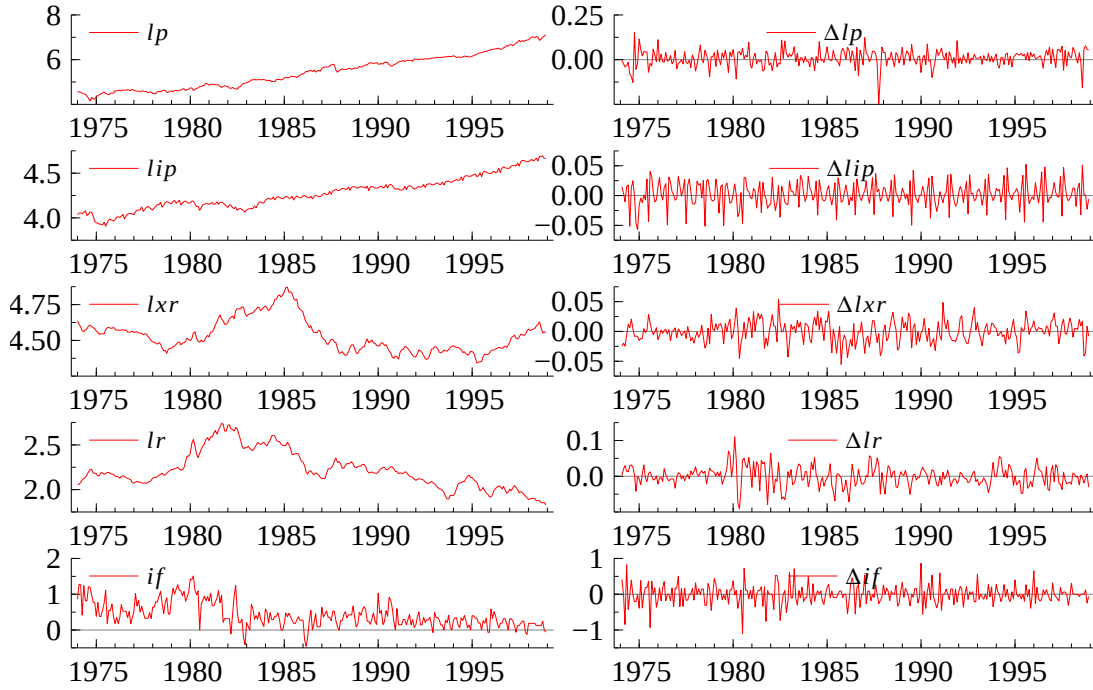


Figure 1: Data in levels and first differences.

6. An application to monthly US data

This section summarizes the application reported in Paruolo (2003), to which we refer for full details. In the analysis, a nominal size of 1% was employed in each test. The data consists of the US monthly time series data analyzed in Kim (2003): the log of S&P 500 composite stock price index, lp , the log of index of industrial production, lip , the log of the US dollar exchange rate, lxr , the log of Moody's Aaa corporate bond yield, lr , and the inflation rate if , calculated as the first difference of the log of the CPI index times 100. The exchange rate series lxr is trade-weighted and price adjusted. The time series are graphed in Fig. 1 in levels and first differences.

These time series were selected by Kim to investigate the relation between the stock price index and its fundamentals, with special reference to exchange rates. I reconsidered the analysis from the choice of lag length with the addition of a constant, seasonal and intervention dummies.⁸ For lag length $k = 13$ the standard mis-specification tests could not reject the hypothesis that the residuals were Gaussian white noise. A preliminary specification analysis was next performed using a LR test on the significance on single variable $\Delta X_{i,t-j}$ in all equations simultaneously; this allowed to drop insignificant variables, reducing the number of variables in V_t from 60 to 15. The resulting estimated residuals are graphed in Fig. 2.

A common trends analysis was next performed. I found one cointegration vector as in Kim (2003), as shown in the Table 1. The cointegration rank was next fixed to $r = 1$. I

⁸The intervention dummies were for the following time periods: 1978:4, 1980:7, 1987:10, 1991:3, 1998:8. ip had a seasonal behavior, and hence seasonal dummies were also included.

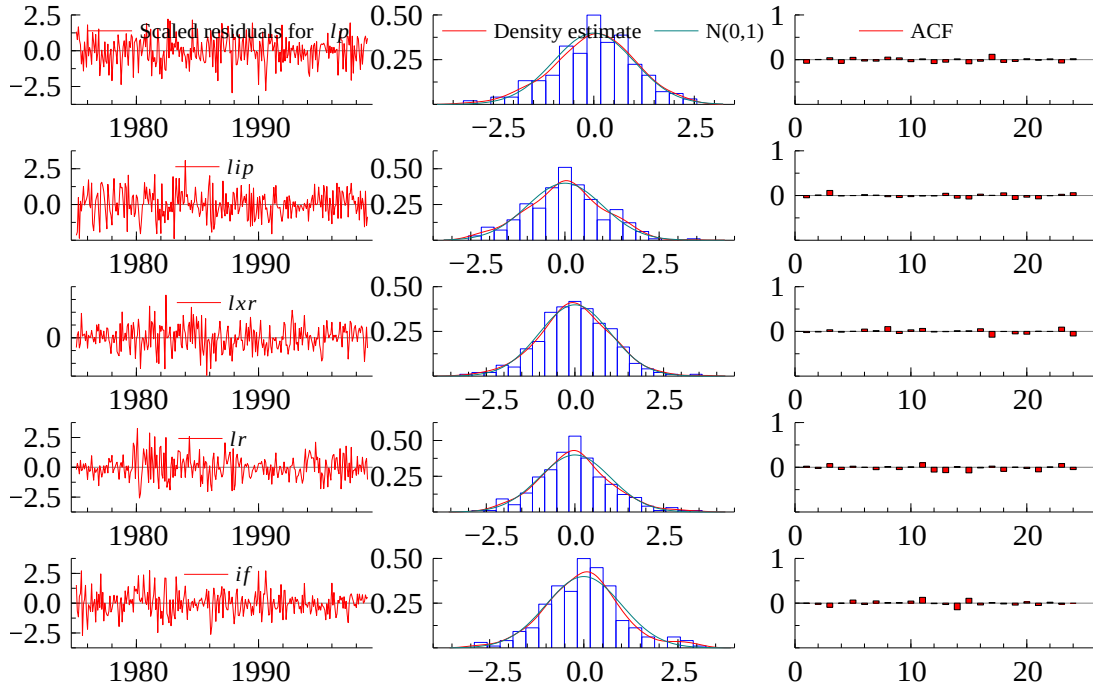


Figure 2: Estimated residuals, their histograms and estimated densities, and their estimated autocorrelation functions, ACF.

H_0	$r = 0$	$r \leq 1$	$r \leq 2$	$r \leq 3$	$r \leq 4$
LR trace test	126.86	52.82	23.62	7.90	2.28
[1% critical values]	[76.37]	[53.91]	[34.87]	[19.69]	[6.64]

Table 1: LR trace test on the cointegration rank r .

next tested the hypothesis

$$H_1 : \beta = \beta_1, \quad \beta_1 := \begin{pmatrix} 1 & -5/2 & 1 & 0 & 3/2 \end{pmatrix}'; \quad (16)$$

the LR test gave a statistic of 3.7368, which compared with a $\chi^2(4)$ variate implies a p -value of 0.4428. This restriction was then imposed, and the transformation $Y_t := (X_t' \beta_1 : \Delta X_t' \beta_{1\perp})'$ was calculated. The transformed time series are defined and graphed in Fig. 3.

I finally performed the cc-cofeature analysis; the results of the LR test $Q_1(s)$ on the cc space dimension are reported in Table 2 both for $Z_{0t} = \Delta X_t$ and $Z_{0t} = Y_t$.

It can be seen that, at the nominal 1% level, one concludes that no cc-cofeature vectors exist for the EqCM form, while $\ell = 1$ cofeature vector exists for the EqD form. In other words, the transformed system presents 4 common cycles in the EqD and one innovation process. The estimated cofeature vector is reported in the Table 3, along with the corresponding asymptotic standard errors. Graphs of the time series $\hat{\varphi}'_{\perp} Y_t$, its histogram and estimated autocorrelation functions are reported in Fig. 4.

7. Conclusions

In this paper two simple forms of common features in VAR processes have been reviewed, each of which leads to one rank restriction on a linear combination of the AR coefficient

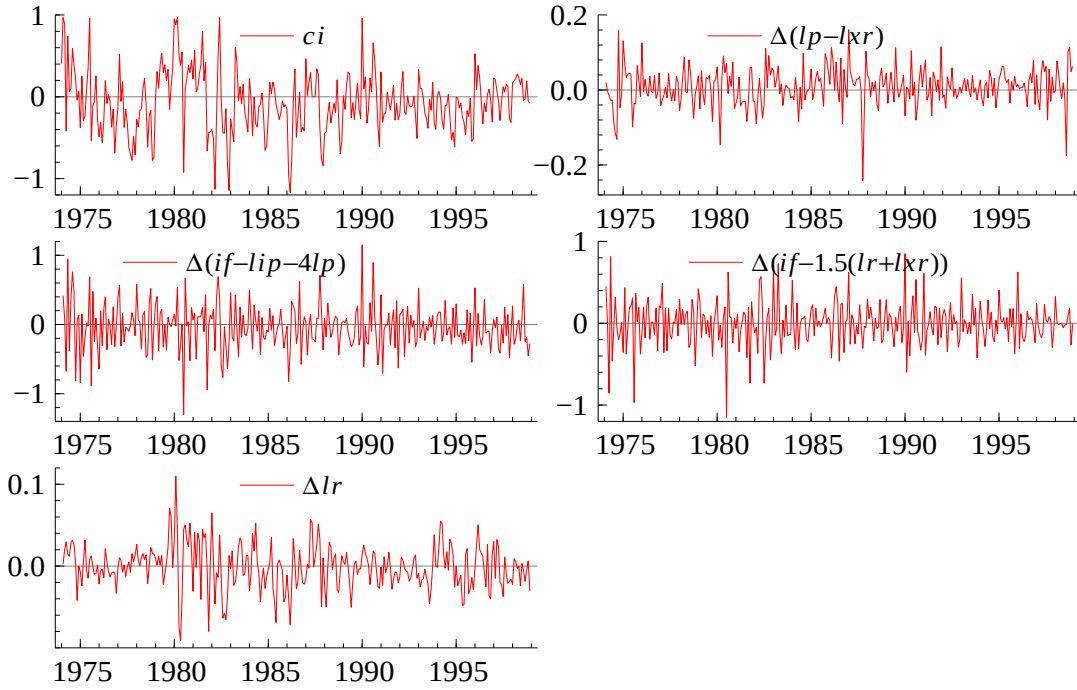


Figure 3: Transformed $Y_t := (X_t' \beta_1 : \Delta X_t' \beta_{1\perp})'$ system.

$H(s)$		$Z_{0t} = \Delta X_t$			$Z_{0t} = Y_t$		
ℓ	s	$Q_1(s)$	df	$100 \cdot p\text{-value}$	$Q_1(s)$	df	$100 \cdot p\text{-value}$
$\ell = 5$	$s = 0$	419.5565	80	0.00	1651.1291	80	0.00
$\ell \geq 4$	$s \leq 1$	248.0962	60	0.00	250.7385	60	0.00
$\ell \geq 3$	$s \leq 2$	144.5994	42	0.00	158.1744	42	0.00
$\ell \geq 2$	$s \leq 3$	67.0296	26	0.00	81.9780	26	0.00
$\ell \geq 1$	$s \leq 4$	26.7864	12	0.83	23.8917	12	2.10

Table 2: LR trace test on the cc dimension s .

matrices. Inference on the dimension of cofeature matrices can be addressed using reduced rank regression, which also delivers ML estimates. The same technique can be used to test various restrictions on the row and column spaces of the reduced rank matrix, as well as hypotheses on bases of their orthogonal complements.

These types of cofeature analysis can be combined into a sequential procedure which starts by analyzing the common trends. It then proceeds by calculating the Y_t transformation using the estimated β and $\beta_{\perp}$ matrices. The subsequent cc analysis is not affected asymptotically by the pre-estimation of β and $\beta_{\perp}$ matrices, because these parameters are superconsistently estimated.

All this provides a useful toolkit for modelling common trends and cycles. One can enquire which variables enter cofeature relations. Economic theory prescribes many equilibrium relations among integrated variables; these are candidate cointegration relations. Other rational expectation models, like the present value models in Campbell and Shiller (1987), imply that certain linear combinations of stationary variables are unpredictable processes. These are candidate cc-cofeature relations. This association between an economic concept – like equilibrium – and a statistical relation – like cointegration – has been

Y'_t	ci	$\Delta(lp - lxr)$	$\Delta(if - lip - 4lp)$	$\Delta(if - \frac{3}{2}(lxx + lr))$	Δlr
$\hat{\varphi}_\perp$	-0.0278	-1.5017	-0.6116	0.5490	1
SE	0.0129	0.4225	0.1235	0.1131	.
t statistics	-2.1563	-3.5539	-4.9527	4.8553	.

Table 3: Estimates of the cc-cofeature vector.

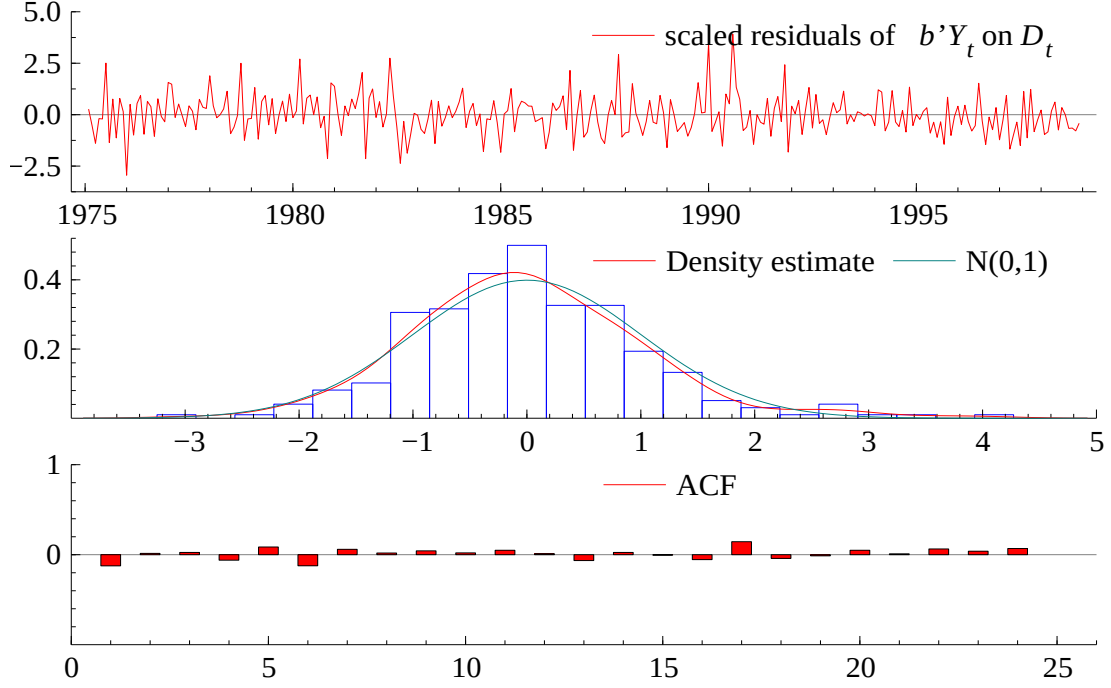


Figure 4: Unpredictable component (top), its histogram and density estimate (middle), estimated autocorrelation function (bottom). D_t contains the deterministic seasonal and intervention dummies, $b = \hat{\varphi}_\perp$.

at the root of the success encountered by cointegration in the last two decades. Hopefully, the just-apposition of cc-cofeatures and rational expectation models could be just as successful.

Applying some of the proposed techniques to the monthly US data analyzed in Kim (2003), one cointegrating relation was found and one cc-cofeature vector was detected; the latter involved the cointegrating relation, and hence could not be found analyzing only growth rates, as it is done in the analysis of the equilibrium correction form.

References

- Ahn S.K. and Reinsel G.C. (1988) Nested reduced rank autoregressive models for multiple time series, *Journal of the American Statistical Association*, 83, 849–856.
- Anderson T.W. (1951) Estimating linear restrictions on regression coefficients for multivariate normal distributions., *Annals of Mathematical Statistics*, 22, 327–351, Correction, *Annals of Statistics* 8, 1980, 1400.
- Campbell J.Y. and Shiller R.J. (1987) Cointegration and tests of present value models, *Journal of Political Economy*, 95, 1062–1087.
- Cubadda G. (1999) Common cycles in seasonal non-stationary time series, *Journal of Applied Econometrics*, 14, 273–291.
- Cubadda G. (2001a) Common features in time series with both deterministic and stochastic seasonality, *Econometric Reviews*, 20, 201–216.
- Cubadda G. (2001b) Complex reduced rank models for seasonally cointegrated time series, *Oxford Bulletin of Economics and Statistics*, 63, 497–511.
- Cubadda G. and Hecq A. (2001) On non-contemporaneous short-run comovements, *Economics Letters*, 73, 389–397.
- Engle R.F. and Granger C.W.J. (1987) Co-integration and error correction: representation, estimation and testing., *Econometrica*, 55, 251–276.
- Engle R.F. and Kozicki S. (1993) Testing for common features, *Journal of Business and Economic Statistics*, 11, 369–380.
- Gonzalo J. and Granger C. (1995) Estimation of common long-memory components in cointegrated systems, *Journal of Business and Economic Statistics*, 13, 27–36.
- Haldrup N. (1998) An econometric analysis of $I(2)$ variables, *Journal of Economic Surveys*, 12, 595–650.
- Hecq A., Palm F. and Urbain J.P. (2002) Separation, weak exogeneity and p-t decompositions in cointegrated var systems with common features, *Econometric Reviews*, 21, 273–307.
- Hendry D. (1995) *Dynamic econometrics*, Oxford: Oxford University Press.
- Johansen S. (1988) Statistical analysis of cointegration vectors, *Journal of Economics Dynamics and Control*, 12, 231–254.
- Johansen S. (1992) Determination of cointegration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics*, 54, 383–397.
- Johansen S. (1996) *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford: Oxford University Press, Second corrected printing.
- Kim K. (2003) Dollar exchange rate and stock price: evidence from a multivariate cointegration and error correction model, *Review of Financial Economics*, 12, 301–313.
- Paruolo P. (2002) Common trend and cycles in $I(2)$ var systems, *WP University of Insubria*, 2002/29.
- Paruolo P. (2003) Common dynamics in $I(1)$ var systems, *WP University of Insubria*, 2003/33.
- Tiao G.C. and Tsay R.S. (1989) Model specification in multivariate time series, *Journal of the Royal Statistical Society, Series B*, 51, 157–213.
- Vahid F. and Engle R. (1993) Common trends and common cycles, *Journal of Applied Econometrics*, 8, 341–360.