A finite, empirically useless and almost sure VAR representation for all minimal transition equations

Saccal, Alessandro

22 February 2023

A finite, empirically useless and almost sure VAR representation for all minimal transition equations

Alessandro Saccal*

February 22, 2023

> *"... the probabilistically surest empirical representation of first order... LRE models and of DSGE ones in particular is not merely a linear combination of non-minimal states in principle, whenever not equal to zero, but no more than a pseudo-variance shifted vector of white noises for which no kind of forecast analysis is at all conceivable."*

Abstract

Does there exist a systematic manner to derive a finite vector autoregression (VAR) representation for any minimal transition equation? While the good news be that any transition equation of a minimal linear time invariant (LTI) state space representation in discrete time admits a VAR representation of finite order of the non-minimal states in the (minimal) measurement equation's outputs, the bad news are that such a representation, on account of the procedure underlying its derivation, is both the probabilistically surest and empirically useless, ranging from linear combinations of non-minimal states in principle, equal to shifted white noises, to output nullity, thereby presenting negative repercussions with particular regard to first order linear rational expectations (LRE) models of optimising representative agents.

JEL classification codes: C02; C32.
MSC codes: 91B51; 93B20.
Keywords: DSGE models; LRE models; minimality; state space; VMA representation; VAR representation.

1. Introduction

This work's methodological contribution is the presentation of a systematic manner to derive a finite VAR representation for any minimal transition equation in discrete time. This work's notional contribution is nevertheless the clarification that such a representation, on account of the procedure underlying its derivation, is both the probabilistically surest and empirically useless, ranging from linear combinations of non-minimal states in principle, equal to shifted white noises, to output nullity, thereby presenting negative repercussions with particular regard to first order LRE models and dynamic stochastic general equilibrium (DSGE) models.

Sections 2 and 3 introduce state space and VAR representations as well as minimality. Findings are adduced in section 4. Section 5 offers applications. Section 6 concludes.

2. State space and VAR representations

*saccal.alessandro@gmail.com. Disclaimer: the author has no declaration of interest related to this research; all views and errors in this research are the author's.

Consider an LTI state space representation in discrete time, in which x_t is a vector of states (i.e. endogenous variables), u_t is a vector of inputs or controls (i.e. exogenous shocks) and y_t is a vector of outputs or observables:

$$\begin{aligned} x_t &= Ax_{t-1} + Bu_t \\ y_t &= Cx_{t-1} + Du_t, \end{aligned}$$

$\forall t \in \mathbb{Z},\ x_t \in \mathbb{R}^{n_x},\ u_t \in \mathbb{R}^{n_u},\ y_t \in \mathbb{R}^{n_y},\ A \in \mathbb{R}^{n_x \times n_x},\ B \in \mathbb{R}^{n_x \times n_u},\ M \in \mathbb{R}^{n_y \times n_x},\ C \in \mathbb{R}^{n_y \times n_x}$ and $D \in \mathbb{R}^{n_y \times n_u}$, observing that $Mx_t = MAx_{t-1} + MBu_t \longleftrightarrow y_t = Cx_{t-1} + Du_t$.

The first equation is termed the transition or state equation, M is termed the measurement or observation matrix, which is normally a selector matrix (i.e. of all zeros and a single one per row), and the second equation is termed the measurement or observation equation.

Let D be invertible and therefore square: D^{-1} such that $n_y = n_u \leq n_x \in \mathbb{N}_+$, (i) either adding $n_y - n_u$ measurement errors to u_t or dropping $n_y - n_u$ from y_t whenever $n_y > n_u$ and (ii) dropping inputs whenever $n_y < n_u$. Solve then for u_t in the measurement equation and plug it into the transition equation:

$$\begin{aligned} & y_t = Cx_{t-1} + Du_t \longrightarrow \\ & \longrightarrow u_t = D^{-1}\left(y_t - Cx_{t-1}\right) \longrightarrow \\ & \longrightarrow x_t = Ax_{t-1} + BD^{-1}\left(y_t - Cx_{t-1}\right) \longrightarrow \\ & \longrightarrow x_t = \left(A - BD^{-1}C\right)x_{t-1} + BD^{-1}y_t = Fx_{t-1} + BD^{-1}y_t, \end{aligned}$$

observing that matrix $F = A - BD^{-1}C$. Solve equation $x_t = Fx_{t-1} + BD^{-1}y_t$ backwards and plug it therefrom into the measurement equation:

$$\begin{aligned} & x_t = Fx_{t-1} + BD^{-1}y_t \longrightarrow \\ & \longrightarrow (I - FL)\,x_t = BD^{-1}y_t \longrightarrow \\ & \longrightarrow x_t = (I - FL)^{-1}\,BD^{-1}y_t \longrightarrow \\ & \longrightarrow x_t = \sum_{j=0}^{\infty} F^j L^j BD^{-1}y_t \longrightarrow \\ & \longrightarrow x_t = \sum_{j=0}^{\infty} F^j BD^{-1}y_{t-j} \longrightarrow \\ & \longrightarrow y_t = Cx_{t-1} + Du_t = C\sum_{j=0}^{\infty} F^j BD^{-1}y_{t-j-1} + Du_t, \end{aligned}$$

in which operator $L = x_t^{-1}x_{t-1}$. Observe that equation $x_t = \sum_{j=0}^{\infty} F^j BD^{-1}y_{t-j}$ is causal and that matrix $(I - FL)^{-1} = \sum_{j=0}^{\infty} F^j L^j$ if and only if F's characteristic polynomial eigenvalues are less than one in modulus (i.e. F is stable), that is, F's elements are square summable (i.e. F's trace is finite).

Formally: $x_t = \sum_{j=0}^{\infty} F^j BD^{-1}y_{t-j}$ is causal and $(I - FL)^{-1} = \sum_{j=0}^{\infty} F^j L^j \longleftrightarrow I = (I - FL)\sum_{j=0}^{\infty} F^j L^j$ if and only if $|\lambda_{F(\lambda)}| < 1$ for $F(\lambda) = F - \lambda I$ in $det\,[F(\lambda)] = 0$, that is, $tr\left(FF^\top\right) = \sum_{i,\ j=1}^{n_x} f_{ij} f_{ji}^\top = \sum_{i,\ j=1}^{n_x} f_{ij}^2 < \infty$.

3. VAR representations and minimality

Accordingly, Fernández-Villaverde *et al.* [5] prove that equation $y_t = C\sum_{j=0}^{\infty} F^j BD^{-1}y_{t-j-1} + Du_t$ is a vector autoregression of infinite order $VAR(\infty)$ for $u_t \sim \mathcal{N}\left(0,\ \Sigma_u^2\right)$ and thereby fundamental if F is stable, terming it the poor man's invertibility condition (PMIC). In other words, if F's trace is finite there then exists a $VAR(\infty)$ of x_t in y_t.

Ravenna [13] and Franchi and Vidotto [8] prove that if F's characteristic polynomial eigenvalues are zero (i.e. F is nilpotent) there then exists a vector autoregression of finite order $VAR(k)$ for $k \in \mathbb{N}_+$ of x_t in y_t : $y_t = C\sum_{j=0}^{k} F^j BD^{-1}y_{t-j-1} + Du_t$ if and only if $\lambda_{F(\lambda)} = 0$.

Franchi [6] and Franchi and Paruolo [7] observe that F's stability is merely sufficient, but unnecessary, for a $VAR(\infty)$ of x_t in y_t, unlike the stability of a minimal F, being sufficient and necessary thereby: F_m is stable if and only if there exists $VAR(\infty)$ of x_t in y_t, which one terms the minimal poor man's minimal invertibility condition (mPMIC) hereby. The reason is that while a stable F may not originate a $VAR(\infty)$ of x_t in y_t a stable F_m could.

Crucially, Franchi [6] and Franchi and Paruolo [7] observe that in minimal LTI state space representations the impulse response functions (IRFs) of the transition equation and the coefficients of the VAR and moving average (MA) representations of x_t in y_t are invariant: $\forall j \in \mathbb{N}_+,\ CA^jB = C_mA_m^jB_m \neq 0$, from

$$
\begin{aligned}
& x_t = Ax_{t-1} + Bu_t \longrightarrow \\
& \longrightarrow (I - AL)\,x_t = Bu_t \longrightarrow \\
& \longrightarrow x_t = (I - AL)^{-1}\,Bu_t \longrightarrow \\
& \longrightarrow x_t = \sum_{j=0}^{\infty} A^jL^jBu_t \longrightarrow \\
& \longrightarrow x_t = \sum_{j=0}^{\infty} A^jBu_{t-j} \longrightarrow \\
& \longrightarrow y_t = Cx_{t-1} + Du_t = C\sum_{j=0}^{\infty} A^jBu_{t-j} + Du_t
\end{aligned}
$$

if and only if $|\lambda_{A(\lambda)}| \gtreqless 1$ for $A(\lambda) = A - \lambda I$ in $det\,[A(\lambda)] = 0$, and $CF^jB = C_mF_m^jB_m \neq 0$, from $y_t = C_m\sum_{j=0}^{\infty} F_m^jB_mD^{-1}y_{t-j-1} + Du_t$.

LTI state space representations are minimal if and only if rank $n_x = r_{\mathcal{C}} = r_{\mathcal{O}}$ for controllability matrix $\mathcal{C} = \left[B \cdots\ A^{n_x-1}B\right]$ and observability matrix $\mathcal{O} = \left[C \cdots\ CA^{n_x-1}\right]^\top$. Non-minimal representations can be reduced to minimal ones by the Kalman decomposition, practically deployed by means of the three following steps.

Step 1. Construct $\mathcal{C} = \left[B \cdots\ A^{n_x-1}B\right]$ and $\mathcal{O} = \left[C \cdots\ CA^{n_x-1}\right]^\top$.

Step 2. If $n_x = r_{\mathcal{C}}$ then the representation is controllable: $\bar{x}_{ct},\ \bar{A}_c,\ \bar{B}_c,\ \bar{C}_c,\ \bar{\mathcal{C}}_c$ and $\bar{\mathcal{O}}_c$; go to Step 3. If $n_x > r_{\mathcal{C}}$ then construct similarity transformation matrix $\mathcal{T} = [\mathcal{C}_{r_{\mathcal{C}}}\ v_{n_x - r_{\mathcal{C}}}]$ such that $\bar{x}_{c\bar{c}t} = \mathcal{T}^{-1}x_t,\ \bar{A}_{c\bar{c}} = \mathcal{T}^{-1}A\mathcal{T},\ \bar{B}_{c\bar{c}} = \mathcal{T}^{-1}B,\ \bar{C}_{c\bar{c}} = C\mathcal{T},\ \bar{\mathcal{C}}_{c\bar{c}} = \mathcal{T}^{-1}\mathcal{C}$ and $\bar{\mathcal{O}}_{c\bar{c}} = \mathcal{O}\mathcal{T}$. The representation is controllable in the first $r_{\mathcal{C}}$ states: $\bar{x}_{ct},\ \bar{A}_c,\ \bar{B}_c,\ \bar{C}_c,\ \bar{\mathcal{C}}_c$ and $\bar{\mathcal{O}}_c$; go to Step 3.

Step 3. If $n_{\bar{x}_c} = r_{\bar{\mathcal{O}}_c}$ then the representation is controllable and observable (i.e. minimal): $\bar{x}_{cot} = x_{mt},\ \bar{A}_{co} = A_m,\ \bar{B}_{co} = B_m,\ \bar{C}_{co} = C_m,\ \bar{\mathcal{C}}_{co} = \mathcal{C}_m$ and $\bar{\mathcal{O}}_{co} = \mathcal{O}_m$. If $n_{\bar{x}_c} > r_{\bar{\mathcal{O}}_c}$ then construct similarity transformation matrix $\mathcal{T} = \left[\bar{\mathcal{O}}_{cr_{\bar{\mathcal{O}}_c}}\ v_{n_{\bar{x}_c} - r_{\bar{\mathcal{O}}_c}}\right]^\top$ such that $\bar{x}_{co\bar{o}t} = \mathcal{T}^{-1}\bar{x}_{ct},\ \bar{A}_{co\bar{o}} = \mathcal{T}^{-1}\bar{A}_c\mathcal{T},\ \bar{B}_{co\bar{o}} = \mathcal{T}^{-1}\bar{B}_c,\ \bar{C}_{co\bar{o}} = \bar{C}_c\mathcal{T},\ \bar{\mathcal{C}}_{co\bar{o}} = \mathcal{T}^{-1}\bar{\mathcal{C}}_c$ and $\bar{\mathcal{O}}_{co\bar{o}} = \bar{\mathcal{O}}_c\mathcal{T}$. The representation is controllable and observable (i.e. minimal) in the first $r_{\bar{\mathcal{O}}_c}$ states: $\bar{x}_{cot} = x_{mt},\ \bar{A}_{co} = A_m,\ \bar{B}_{co} = B_m,\ \bar{C}_{co} = C_m,\ \bar{\mathcal{C}}_{co} = \mathcal{C}_m$ and $\bar{\mathcal{O}}_{co} = \mathcal{O}_m$.

Observe that the reduction of the LTI state space representation to observability before controllability, whenever applicable, does not substantially alter the algorithm: the order of reduction is accidental. All else equal, a minimal LTI state space representation in discrete time is thus the following:

$$
\begin{aligned}
x_{mt} &= A_mx_{mt-1} + B_mu_t \\
y_t &= C_mx_{mt-1} + Du_t,
\end{aligned}
$$

$\forall x_{mt} \in \mathbb{R}^{n_{x_m}}$, $A_m \in \mathbb{R}^{n_{x_m} \times n_{x_m}}$, $B_m \in \mathbb{R}^{n_{x_m} \times n_u}$ and $C_m \in \mathbb{R}^{n_y \times n_{x_m}}$, whereby $n_{x_m} = r_{\mathcal{C}_m} = r_{\mathcal{O}_m}$.

4. $VAR(k)$ representation ubiquity

4.1 $\mathbf{C = 0}$. Consider the case of dimension $n_y = n_u \leq n_x \in \mathbb{N}_+$ and $C = 0$. It follows that $n_x > r_{\mathcal{O}} = \mathcal{O} = \left[C \cdots \; CA^{n_x - 1}\right]^\top = 0$. Consequently, the minimal LTI state space representation is:

$$0$$

$$y_t = Du_t,$$

whereby $n_{x_m} = r_{\mathcal{C}_m} = r_{\mathcal{O}_m} = 0$, so that $x_{mt} = A_m - B_m D^{-1} C_m = F_m = \lambda_{F_m(\lambda)} = 0$. Such signifies that there exists a $VAR(k)$ representation of x_t in y_t, being precisely $y_t = Du_t$. Observe the non-necessity of F's stability for a VAR representation of x_t in y_t : $F = A - BD^{-1}C = A$ and $|\lambda_{A(\lambda)}| \gtreqless 1$ for $A(\lambda) = A - \lambda I$ in $det\left[A(\lambda)\right] = 0$.

The central question then is: can $C = 0$ be categorically guaranteed? Does there exist a systematic manner to obtain $C = 0$ to the end of deriving a finite VAR representation for any minimal transition equation? Even though the answer be a yes its content is not as appealing as it may sound.

4.2 Transpose kernel of $\mathbf{A}^\top$. Provided the existence of a non-trivial kernel or null space of $A^\top$ to begin with, one can systematically obtain $C = 0$ by choosing an $M \neq 0$ such that $MA = C = 0$ in which the non-trivial rows of M differ and are orthogonal to all columns of A, thereby belonging to the transpose kernel of $A^\top$: $\forall h = 1, \ldots, n_y$, $i = 1, \ldots, n_x$, $j = 1, \ldots, n_x$ and $k = 1, \ldots, n_x$, $m_{hi} a_{jk} \neq m_{\neg hi} a_{jk} = 0$, in which $0 \neq M \in ker^\top\left(A^\top\right) = \left\{v^\top \in \mathbb{R}^{n_y \times n_x} : v^\top A = 0\right\}$ and $ker\left(A^\top\right) = \left\{v \in \mathbb{R}^{n_x \times n_y} : A^\top v = 0\right\}$, since $A^\top v = 0 \in \mathbb{R}^{n_x \times n_y} \longrightarrow \left(A^\top v\right)^\top = v^\top A = 0^\top = 0 \in \mathbb{R}^{n_y \times n_x}$, observing that $M = v^\top$ whenever $v^\top \neq 0$.

Basis n_y of the non-trivial kernel of $A^\top$, $v^\top \neq 0$, could certainly exceed one. If such a basis does not exceed one then M is not a matrix, but a non-trivial row vector, and selects one state at best, that is, $n_y = 1$ for $0 \neq M = v^\top \in \mathbb{R}^{1 \times n_x}$, implying $n_u = 1$ and thereby restricting the input potential of both the finite VAR representation and the underlying structural model.

The reason for which M would thereby be a non-trivial row vector, rather than a matrix, is not so much due to its equation with the non-trivial transpose kernel of $A^\top$, $v^\top \neq 0$, as that its artificial rows would otherwise be identical and thereby give rise to a $VAR(k)$ of x_t in y_t by the identical variables, that is, to a y_t vector by the identical entries.

4.3 $\mathbf{n_u = 1}$ restriction. The $n_u = 1$ restriction of the input potential of both the finite VAR representation and the underlying structural model is noticeable. Constraining empirical VARs to the presentation of one exogenous shock alone effectively signifies renouncing to no less than the computation of IRFs and forecast error variance decomposition (FEVDs) in relation to variations in more than one endogenous variable, but the presence of one exogenous shock alone stems precisely from that of one endogenous variable (i.e. a single state at best).

In the underlying structural models, such as DSGE models, for instance, $n_u = 1$ would similarly require the presence of one sole exogenous shock. At first order approximations both the methodological loss and the notional loss could nonetheless be negligible: at a methodological level if the variances proper to more than one exogenous shock, relative its own process, were normalised to the same value then there would effectively exist one sole exogenous shock all the same, irrespective of the shape of the exogenous shock vector's matrix; at a notional level a sole exogenous shock driving variations in more than one process reflects the idea of a single, underlying unexpected variation.

4.4 Vector moving average of order zero $[\mathbf{VMA}(\mathbf{0})]$. The minimal LTI state space representation originating from the exploitation of the non-trivial transpose kernel of $A^\top$, $v^\top \neq 0$, is therefore that shown above, that is, 0 and $y_t = Du_t$, for which $\lambda_{F_m(\lambda)} = 0$.

Such signifies that by said construction of $C = 0$ there categorically exists an accompanying $VAR(k)$ of x_t in y_t, being none other than $y_t = Du_t$ and in fact admitting an $AR(k)$ as well, owing to the fact that the non-trivial transpose kernel of $A^\top$, $v^\top \neq 0$, could be a row vector.

To be even more precise, $y_t = Du_t$ is not even a $VAR(k)$, but a $VMA(0)$, that is, a D pseudo-variance shifted vector of white noises: $y_t = Du_t$, in which $u_t \sim \mathcal{N}(0,\ \Sigma_u^2)$, in which $\Sigma_u \neq D = MB$ and Σ_u need not equal B.

Moreover, the non-trivial transpose kernel of $A^\top,\ v^\top \neq 0$, is not necessarily a selector matrix, being composed of all zeros and a single one per row, but effectively depends on A's elements; because of such a dependence, in fact, it almost never is, arising from linear structural models whose coefficients, even after the Gauss Jordan elimination required for the kernel of $A^\top$'s computation, render it more elaborate.

As alluded to, it could even result that a non-trivial transpose kernel of $A^\top,\ v^\top \neq 0$, may not exist, that is, it could even result that the transpose kernel of $A^\top$ be no more than trivial, $A^\top$ and A being thereby invertible, whereby $M = v^\top = ker^\top(A^\top) = 0$, so that $M = 0$ and $Mx_t = y_t = 0$, being there no outputs at all.

If the non-trivial transpose kernel of $A^\top,\ v^\top \neq 0$, is not a selector matrix then the empirical interpretation of the accompanying $VMA(0)$ of x_t in y_t is such that IRFs and other forecasting tools, as moments and FEVDs, concern not single non-minimal states, but effective linear combinations thereof (e.g. $-2x_{1t} + x_{2t}$ for $n_y = 1$ as in Example 5.1 below), losses in forecasting expedience being therefore substantial.

Indeed, if the basis of the non-trivial kernel of $A^\top,\ v^\top \neq 0$, exceeded one such that M were not a selector matrix then the accompanying $VAR(k)$ representation of x_t in y_t would emerge as even more convoluted than if said basis equalled one.

Regardless of whether one were able to systematically restrict the computation of the non-trivial transpose kernel of $A^\top,\ v^\top \neq 0$, to single non-minimal states, M being a selector matrix thereby, provided a multiple basis of the transpose kernel of $A^\top$ to start with, a linear combination of non-minimal states in principle emerges as a categorical $VAR(k)$ representation of x_t in y_t, as well as the probabilistically surest.

4.5 VMA $(\mathbf{0})$ sterility. Observe that for $x_{mt} = 0$ the IRFs of the transition equation and the coefficients of the VAR and MA representations of x_t in y_t clearly differ from their non-minimal ones: $\forall j \in \mathbb{N}_+,\ CA^jB \neq C_mA_m^jB_m = 0$ and $CF^jB \neq C_mF_m^jB_m = 0$.

Such is not anomalous, however, for the minimal states are zero such that no analysis can be conducted in terms of outputs, that is, no empirical forecasting is possible by definition, be it the computation of IRFs or FEVDs, be it that of moments, be it the recovery of exogenous shocks. Otherwise stated, the nullity of minimal states does not invalidate the ubiquity of a VAR representation of x_t in y_t, but reinforces its inexpedience.

Such an inexpedience precisely consists in $VMA(0)$ representation $y_t = Du_t$'s methodological sterility, which consistently confirms that a linear combination of non-minimal states y_t in principle is tantamount to a pseudo-variance shifted vector of white noises Du_t.

On further reflexion, although, said sterility can be discerned as being effectively insightful, for it does not merely reveal that in empirical terms the underlying structural model (first order DSGE, LRE etc.) is represented at best as a convolution but that such a maximal convolution is itself represented as exogenous shocks and no more, whereby no empirical analysis is in principle possible and whereby the determination of said convolution in principle on the part of exogenous shocks is not merely ultimate but frontal as well.

The last phrase is to mean that exogenous shocks are not merely recovered by means of a finite VAR but through a VMA of order zero, that is, they are recovered almost directly as observables. Otherwise stated, the exogenous shocks of the underlying structural model are the shifted observables themselves.

In sum, while the good news be that any transition equation of a minimal LTI state space representation in discrete time admits a VAR representation of finite order of the non-minimal states in the (minimal) measurement equation's outputs, the bad news are that such a representation, on account of the procedure underlying its derivation, is both the probabilistically surest and empirically useless, ranging from material inutility (i.e. linear combinations of non-minimal states in principle, equal to shifted white noises) to formal inutility (i.e. output nullity). They are also formally adduced by means of the following two propositions.

PROPOSITION 4.6 (Empirical representation ubiquity) *For any minimal transition equation in discrete time there exists a $VMA(0)$ representation of the non-minimal states in the (minimal) measurement equation's outputs (i.e. equal to a pseudo-variance shifted vector of white noises). Additionally, if the minimal transition equation's companion matrix is invertible then said representation equals zero. Formally:* $\forall u_t \sim \mathcal{N}(0,\ \Sigma_u^2)$,

$$\forall x_{mt} = A_m x_{mt-1} + B_m u_t,\ \exists y_t = Du_t;$$
$$\exists A^{-1} \longrightarrow y_t = 0.$$

Proof. Consider an LTI state space representation in discrete time: $x_t = Ax_{t-1} + Bu_t$ and $y_t = Cx_{t-1} + Du_t$, *ceteris paribus.* For any A choose an M such that $MA = C = 0$ in which the rows of M differ and are orthogonal to all columns of A; in other words, choose an M equal to the transpose kernel of $A^\top:\ \forall h = 1,\ \ldots,\ n_y,\ i = 1,\ \ldots,\ n_x,\ j = 1,\ \ldots,\ n_x$ and $k = 1,\ \ldots,\ n_x,\ m_{hi}a_{jk} \neq m_{\neg hi}a_{jk} = 0$, in which $M = ker^\top\left(A^\top\right) = \left\{v^\top \in \mathbb{R}^{n_y \times n_x} : v^\top A = 0\right\}$ and $ker\left(A^\top\right) = \left\{v \in \mathbb{R}^{n_x \times n_y} : A^\top v = 0\right\}$, since $A^\top v = 0 \in \mathbb{R}^{n_x \times n_y} \longrightarrow \left(A^\top v\right)^\top = v^\top A = 0^\top = 0 \in \mathbb{R}^{n_y \times n_x}$, observing that $M = v^\top$.

It follows that the number of non-minimal states exceeds the rank of the observability matrix, which equals zero, consequently, the minimal LTI state space representation is a zero transition equation and a measurement equation equal to a pseudo-variance shifted vector of white noises: $n_x > r_{\mathcal{O}} = \mathcal{O} = \left[C \cdots\ CA^{n_x - 1}\right]^\top = 0$, thus, the minimal LTI state space representation is 0 and $y_t = Du_t$, whereby $n_{x_m} = r_{\mathcal{C}_m} = r_{\mathcal{O}_m} = 0$.

Owing to the satisfaction of the mPMIC, the non-minimal states can be precisely represented as the minimal measurement equation: $x_{mt} = A_m - B_m D^{-1} C_m = F_m = \lambda_{F_m(\lambda)} = 0$, thus, x_t can be represented as $y_t = Du_t$.

Additionally, A's invertibility is necessary and sufficient for that of $A^\top$, signifying that $A^\top v = 0$ admits of one solution alone, being zero, so that the (minimal) measurement equation's outputs be zero: $\exists A^{-1}$ if and only if $\exists \left(A^\top\right)^{-1}$, which signifies that the only solution of $A^\top v = 0$ is $v = 0$, so that $0 = v^\top = M = Mx_t = y_t$.

Observe that A's invertibility is sufficient, but unnecessary for the (minimal) measurement equation's outputs to be zero, since they could also be zero as result of (i) another derivation of $C = \mathcal{O} = 0$ (for $C = MA$ e.g. $M = v^\top \neq 0$ tautologically, $0 \neq M \neq v^\top\big)$, (ii) $C \neq 0$, but $\mathcal{O} = 0$, (iii) $B = \mathcal{C} = 0$, at the cost of $Bu_t = 0$, or (iv) $B \neq 0$, but $\mathcal{C} = 0$, the inclusive disjunction applying wherever suitable. *QED*

Proposition 4.7 (Quasi-surety) *For any minimal transition equation in discrete time a $VMA(0)$ representation of the non-minimal states in the (minimal) measurement equation's outputs is almost surely verified. Formally:* $\forall\left(X_m,\ S,\ \mu\right)$, *in which*

$$X_m = \left\{x_m \in X_m : x_{mt} = A_m x_{mt-1} + B_m u_t\right\},$$
$$\neg P = \left\{x_m \in X_m : \neg P\left(x_m\right)\right\} \subset N \in S \subset \mathcal{P}\left(X_m\right),$$
$$P = N^c = \left\{x_m \in N^c = X_m \backslash N : P\left(x_m\right)\right\} \in S,$$
$$P\left(x_m\right) := \left\{\forall x_{mt},\ \exists! y_t : y_t = Du_t\right\} \textit{ and}$$
$$\neg P\left(x_m\right) := \left\{\forall x_{mt},\ \exists\left\{y_{it}\right\}_{i=1}^n : \left(y_{it} = Du_t\right) \wedge \left(y_{\neg it} = C_m \sum_{j=0}^{\infty} F_m^j B_m D^{-1} y_{\neg it-j-1} + Du_t,\ \forall C_m F_m^j B_m \neq 0\right)\right\},$$
$$\mu\left(N^c\right) = \mu\left(P\right) = 1 \textit{ and } \mu\left(N\right) = 0.$$

Proof. μ can be understood as a probability measure such that $\int_{-\infty}^{\infty} \mu\left(x_m\right) dx_m = 1$ and $\mu\left(x_m\right) dx_m \geq 0$. Now, by Proposition 4.6 $\mu\left(N^c\right) = \mu\left(P\right) = 1$.

Additionally, by the ubiquity of minimal LTI state space representations in discrete time $\mu\left(X_m\right) = 1$. Observe that $\mu\left(X_m\right) = \mu\left(N\right) + \mu\left(N^c\right)$, since $\mu\left(X_m\right) = 1 < \infty$ (i.e. finite). Therefore, $\mu\left(N\right) = \mu\left(X_m\right) - \mu\left(N^c\right) = \mu\left(X_m\right) - \mu\left(P\right) = 1 - 1 = 0$, whereby $P = N^c$ is a co-null set.

Consequently, a $VMA(0)$ representation of the non-minimal states in the (minimal) measurement equation's outputs occurs almost surely, almost everywhere in X_m or for almost every $x_m \in X_m$, that is, for almost every minimal transition equation, so that those for which it is not verified (i.e. other empirical representations thereof are also verified) are negligible. *QED*

The reason for which the negation of the property evoking a $VMA(0)$ representation of x_t in y_t as univocal is defined as evoking more than one $VAR(\infty)$ representation x_t in y_t is self-evident, that is, axiomatic or indemonstrable. Comparably, all other proper subsets of power set $\mathcal{P}(X_m)$ as well as any subset thereof or even any other element of set S refer to other properties of universe X_m.

5. Applications and LRE models discussion

Example 5.1 (Generic) Consider an LTI state space representation in discrete time and let

$$\begin{aligned} x_t &= \begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix}, \; u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \\ A &= \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} \sigma_{u_1} & 0 \\ 0 & \sigma_{u_2} \\ \sigma_{u_1} & \sigma_{u_2} \end{bmatrix}, \end{aligned}$$

in which $u_t \sim \mathcal{N}\left(0, \; \Sigma_u^2\right)$ and $\sigma_{u_{1,\,2}} \in \mathbb{R}$, more specifically. The transition equation is thus

$$\begin{aligned} & x_t = Ax_{t-1} + Bu_t \longleftrightarrow \\ & \longleftrightarrow \begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} + \begin{bmatrix} \sigma_{u_1} & 0 \\ 0 & \sigma_{u_2} \\ \sigma_{u_1} & \sigma_{u_2} \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \end{aligned}$$

Compute the non-trivial transpose kernel of $A^\top$, $v^\top \neq 0$:

$$\begin{aligned} A^\top &= \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix} \longrightarrow \\ &\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 & 1 & 0 \\ 3 & 6 & 1 & 0 & 0 & 1 \end{array}\right] \longrightarrow \\ &\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 & 0 \\ 3 & 6 & 1 & 0 & 0 & 1 \end{array}\right] \tilde{r}_2 = r_2 - 2r_1 \end{aligned}$$

in which $M = v^\top = \begin{bmatrix} -2 & 1 & 0 \end{bmatrix} \in ker^\top\left(A^\top\right)$, being the sole echelon form (non-zero) row of the identity sub-matrix corresponding to an echelon form zero row of $A^\top$. It follows that

$$\begin{aligned} & Mx_t = y_t \longleftrightarrow \\ & \longleftrightarrow \begin{bmatrix} -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} = -2x_{1t} + x_{2t}. \end{aligned}$$

Consequently, the minimal LTI state space representation guaranteeing a non-zero $VMA(0)$ of x_t in y_t, being an $MA(0)$ hereby, is

$$\begin{aligned} & y_t = MBu_t = Du_t \longleftrightarrow \\ & \longleftrightarrow -2x_{1t} + x_{2t} = \begin{bmatrix} -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{u_i} \\ \sigma_{u_i} \\ 2\sigma_{u_i} \end{bmatrix} u_{it} = -\sigma_{u_i} u_{it}, \end{aligned}$$

for any $i = 1,\ 2,$ having removed either input amongst u_{1t} and u_{2t} from u_t and normalised $\sigma_{u_1} = \sigma_{u_2}$ such that $B = [\sigma_{u_i}\ \sigma_{u_i}\ 2\sigma_{u_i}]^\top$.

Example 5.2 (Hybrid New Keynesian Phillips curve) Consider the vector representation of a hybrid New Keynesian Phillips curve of the linear form[1] $\pi_t = \alpha \mathbb{E}_t \pi_{t+1} + \beta \pi_{t-1} + \gamma \varepsilon_t$, in which $\alpha = (1-a)\, b,\ \beta = a,\ \gamma = (1-h)^{-1} \{h [1-(1-h)\, b]\, c\}$ and $\Theta = \{a,\ b,\ c,\ h\} \subset \mathbb{R}$:

$$\begin{aligned}
& Qx_t = Rx_{t-1} + S\varepsilon_t + T\eta_t \longleftrightarrow \\
& \longleftrightarrow \begin{bmatrix} 1 & -\alpha \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_t \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} \beta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \mathbb{E}_{t-1} \pi_t \end{bmatrix} + \begin{bmatrix} \gamma \\ 0 \end{bmatrix} \varepsilon_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_t,
\end{aligned}$$

in which $\varepsilon_t \sim \mathcal{N}\left(0,\ \sigma_\varepsilon^2\right)$ and $\eta_t = \pi_t - \mathbb{E}_{t-1}\pi_t$ is an expectational error. For simplicity, parametrise $\alpha = 0.6,\ \beta = 0.4$ and $\gamma = 0.01$:

$$\begin{bmatrix} 1 & -0.6 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_t \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \mathbb{E}_{t-1} \pi_t \end{bmatrix} + \begin{bmatrix} 0.01 \\ 0 \end{bmatrix} \varepsilon_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_t.$$

Sims' [14] unique solution algorithm for first order LRE models dictates:

$$\begin{aligned}
& HJ_Q K^\top x_t = HJ_R K^\top x_{t-1} + S\varepsilon_t + T\eta_t \longrightarrow \\
& \longrightarrow HJ_Q z_t = HJ_R z_{t-1} + S\varepsilon_t + T\eta_t \longrightarrow \\
& \longrightarrow J_Q z_t = J_R z_{t-1} + H^\top (S\varepsilon_t + T\eta_t) \longleftrightarrow \\
& \longleftrightarrow \begin{bmatrix} J_{Q11} & J_{Q12} \\ 0 & J_{Q22} \end{bmatrix} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} J_{R11} & J_{R12} \\ 0 & J_{R22} \end{bmatrix} \begin{bmatrix} z_{1t-1} \\ z_{2t-1} \end{bmatrix} + \begin{bmatrix} \hat{H}_{11} & \hat{H}_{12} \\ \hat{H}_{21} & \hat{H}_{22} \end{bmatrix} \left\{ \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \varepsilon_t + \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \eta_t \right\} = \\
& = \begin{bmatrix} J_{R11} & J_{R12} \\ 0 & J_{R22} \end{bmatrix} \begin{bmatrix} z_{1t-1} \\ z_{2t-1} \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \varepsilon_t + \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \eta_t \longrightarrow \\
& \longrightarrow J_{Q11} z_{1t} + J_{Q12} z_{2t} = J_{R11} z_{1t-1} + J_{R12} z_{2t-1} + U_1 \varepsilon_t + V_1 \eta_t \text{ and} \\
& J_{Q22} z_{2t} = J_{R22} z_{2t-1} + U_2 \varepsilon_t + V_2 \eta_t \longrightarrow \\
& \longrightarrow J_{R22} z_{2t-1} = J_{Q22} z_{2t} - U_2 \varepsilon_t - V_2 \eta_t \longrightarrow \\
& \longrightarrow z_{2t-1} = J_{R22}^{-1} J_{Q22} z_{2t} - J_{R22}^{-1} U_2 \varepsilon_t - J_{R22}^{-1} V_2 \eta_t \longrightarrow \\
& \longrightarrow z_{2t} = J_{R22}^{-1} J_{Q22} \mathbb{E}_t z_{2t+1} - J_{R22}^{-1} U_2 \mathbb{E}_t \varepsilon_{t+1} - J_{R22}^{-1} V_2 \mathbb{E}_t \eta_{t+1} = \\
& = J_{R22}^{-1} J_{Q22} \left[J_{R22}^{-1} J_{Q22} \mathbb{E}_t z_{2t+2} - J_{R22}^{-1} U_2 \mathbb{E}_t \varepsilon_{t+2} - J_{R22}^{-1} V_2 \mathbb{E}_t \eta_{t+2} \right] - J_{R22}^{-1} U_2 \mathbb{E}_t \varepsilon_{t+1} - J_{R22}^{-1} V_2 \mathbb{E}_t \eta_{t+1} = \\
& = \left(J_{R22}^{-1} J_{Q22} \right)^2 \mathbb{E}_t z_{2t+2} - J_{R22}^{-2} J_{Q22} U_2 \mathbb{E}_t \varepsilon_{t+2} - J_{R22}^{-2} J_{Q22} V_2 \mathbb{E}_t \eta_{t+2} - J_{R22}^{-1} U_2 \mathbb{E}_t \varepsilon_{t+1} - J_{R22}^{-1} V_2 \mathbb{E}_t \eta_{t+1} \longrightarrow \\
& \longrightarrow z_{2t} = \lim_{j \to \infty} \left(J_{R22}^{-1} J_{Q22} \right)^j \mathbb{E}_t z_{2t+j} - \sum_{j=1}^{\infty} J_{R22}^{-j} J_{Q22} U_2 \mathbb{E}_t \varepsilon_{t+j} - \sum_{j=1}^{\infty} J_{R22}^{-j} J_{Q22} V_2 \mathbb{E}_t \eta_{t+j} = 0 \longrightarrow \\
& \longrightarrow U_2 \varepsilon_t + V_2 \eta_t = 0 \longrightarrow \\
& \longrightarrow V_2 \eta_t = -U_2 \varepsilon_t \longrightarrow \\
& \longrightarrow \eta_t = -V_2^{-1} U_2 \varepsilon_t, \text{ provided } n_{\lambda_2} = n_\eta, \text{ and} \\
& J_{Q11} z_{1t} = J_{R11} z_{1t-1} + U_1 \varepsilon_t + V_1 \left(-V_2^{-1} U_2 \varepsilon_t \right) = J_{R11} z_{1t-1} + \left(U_1 - V_1 V_2^{-1} U_2 \right) \varepsilon_t \longrightarrow \\
& \longrightarrow z_{1t} = J_{Q11}^{-1} J_{R11} z_{1t-1} + J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \varepsilon_t, \text{ thus,} \\
& \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} J_{Q11}^{-1} J_{R11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1t-1} \\ z_{2t-1} \end{bmatrix} + \begin{bmatrix} J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \\ 0 \end{bmatrix} \varepsilon_t \longrightarrow
\end{aligned}$$

[1] https://en.wikipedia.org

$$\longrightarrow \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} \pi_{1t} \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} J_{Q11}^{-1} J_{R11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \end{bmatrix} + \begin{bmatrix} J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \\ 0 \end{bmatrix} \varepsilon_t \longrightarrow$$

$$\longrightarrow \begin{bmatrix} \pi_t \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} J_{Q11}^{-1} J_{R11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \mathbb{E}_{t-1} \pi_t \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \\ 0 \end{bmatrix} \varepsilon_t =$$

$$= \begin{bmatrix} J_{Q11}^{-1} J_{R11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \mathbb{E}_{t-1} \pi_t \end{bmatrix} + \begin{bmatrix} K_{11} J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \\ K_{21} J_{Q11}^{-1} \left(U_1 - V_1 V_2^{-1} U_2 \right) \end{bmatrix} \varepsilon_t \longleftrightarrow$$

$$\longleftrightarrow x_t = A x_{t-1} + B \varepsilon_t.$$

More specifically, a generalised Schur decomposition solves the generalised eigenvalue problem $Qv = \lambda Rv$ such that $Q = HJ_QK^\top,\ R = HJ_RK^\top$ and generalised eigenvalue $\lambda_i = \frac{J_{Rii}}{J_{Qii}},\ J_Q$ and J_R eigenvalues being situated along the respective diagonals.

J_Q and J_R are upper triangular and $HH^\top = HH^{-1} = KK^\top = KK^{-1} = I$; in detail, $J_Q,\ J_R \in \mathbb{R}^{n_\lambda \times n_\lambda},\ H,\ K \in \mathbb{R}^{n_x \times n_\lambda}$ and $K^\top,\ H^\top \in \mathbb{R}^{n_\lambda \times n_x}$.

For clarity, $Q,\ R \in \mathbb{R}^{n_x \times n_x},\ S \in \mathbb{R}^{n_x \times n_\varepsilon},\ T \in \mathbb{R}^{n_x \times n_\eta},\ U_1 \in \mathbb{R}^{n_{\lambda_1} \times n_\varepsilon},\ U_2 \in \mathbb{R}^{n_{\lambda_2} \times n_\varepsilon},\ V_1 \in \mathbb{R}^{n_{\lambda_1} \times n_\eta}$ and $V_2 \in \mathbb{R}^{n_{\lambda_2} \times n_\eta}$.

J_Q and J_R are additionally reordered such that J_{Q11} and J_{R11} respectively contain all eigenvalues smaller than one in modulus; accordingly, J_{Q22} and J_{R22} are reordered to contain all eigenvalues no smaller than one in modulus: $|\lambda_{J_Q(\lambda)}| < 1$ in J_{Q11} and $|\lambda_{J_Q(\lambda)}| \geq 1$ in J_{Q22} for $J_Q(\lambda) = J_Q - \lambda I$ in $det\,[J_Q(\lambda)] = 0$; $|\lambda_{J_R(\lambda)}| < 1$ in J_{R11} and $|\lambda_{J_R(\lambda)}| \geq 1$ in J_{R22} for $J_R(\lambda) = J_R - \lambda I$ in $det\,[J_R(\lambda)] = 0$.

Observe that $z_{2t} = 0$ owing to the following facts. Expectational endogenous variable stationarity: $\lim_{j\to\infty} \mathbb{E}_t z_{2t+j} < \infty$. Eigenvalues instability: $\lim_{j\to\infty} J_{R22}^{-j} = 0$. Exogenous shock zero mean (i.e. white noise): $\sum_{j=1}^{\infty} \mathbb{E}_t \varepsilon_{t+j} = 0$. Expectational error zero conditional mean: $\sum_{j=1}^{\infty} \mathbb{E}_t \eta_{t+j} = 0$. The transition equation is thus

$$x_t = A x_{t-1} + B\varepsilon_t \longleftrightarrow$$

$$\longleftrightarrow \begin{bmatrix} \pi_t \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} 0.667 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \mathbb{E}_{t-1} \pi_t \end{bmatrix} + \begin{bmatrix} 0.017 \\ 0.011 \end{bmatrix} \varepsilon_t.$$

Compute the non-trivial transpose kernel of $A^\top,\ v^\top \neq 0$:

$$A^\top = \begin{bmatrix} 0.667 & 0 \\ 0 & 0 \end{bmatrix} \longrightarrow$$

$$\longrightarrow \left[\begin{array}{cc|cc} 0.667 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

in which $M = v^\top = \begin{bmatrix} 0 & 1 \end{bmatrix} \in ker^\top \left(A^\top \right)$, being the sole echelon form (non-zero) row of the identity sub-matrix corresponding to a zero row of $A^\top$. It follows that

$$M x_t = y_t \longleftrightarrow$$

$$\longleftrightarrow \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_t \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \mathbb{E}_t \pi_{t+1}.$$

Consequently, the minimal LTI state space representation guaranteeing a non-zero $VMA\,(0)$ of x_t in y_t, being an $MA\,(0)$ hereby, is

$$y_t = MB\varepsilon_t = D\varepsilon_t \longleftrightarrow$$
$$\longleftrightarrow \mathbb{E}_t\pi_{t+1} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.017 \\ 0.011 \end{bmatrix} \varepsilon_t = 0.011\varepsilon_t,$$

whereby the two non-minimal states π_t and $\mathbb{E}_t\pi_{t+1}$ are empirically represented by the linear combination in principle $m_1\pi_t + m_2\mathbb{E}_t\pi_{t+1} = \mathbb{E}_t\pi_{t+1}$, itself potentially measurable and to be itself represented as pseudo-variance shifted white noise $0.011\varepsilon_t$.

Example 5.3 (Interest rate dynamics) Consider an equation for the nominal interest rate such that the real interest rate and inflation are $AR(1)$ processes (i.e. amnesic) driven by the same exogenous shock: $rn_t = r_t + \mathbb{E}_t\pi_{t+1}$, $r_t = 0.99r_{t-1} + \varepsilon_t$ and $\pi_t = 0.99\pi_{t-1} + \varepsilon_t$, in which $\varepsilon_t \sim \mathcal{N}\left(0,\ \sigma_\varepsilon^2\right)$. The vector representation of such three equations is

$$Qx_t = Rx_{t-1} + S\varepsilon_t + T\eta_t \longleftrightarrow$$
$$\longleftrightarrow \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} rn_t \\ r_t \\ \pi_t \\ \mathbb{E}_t\pi_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.99 & 0 & 0 \\ 0 & 0 & 0.99 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} rn_{t-1} \\ r_{t-1} \\ \pi_{t-1} \\ \mathbb{E}_{t-1}\pi_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \varepsilon_t + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \eta_t,$$

in which $\eta_t = \pi_t - \mathbb{E}_{t-1}\pi_t$. By Sims' [14] unique solution algorithm for first order LRE models, the transition equation is thus

$$x_t = Ax_{t-1} + B\varepsilon_t \longleftrightarrow$$
$$\longleftrightarrow \begin{bmatrix} rn_t \\ r_t \\ \pi_t \\ \mathbb{E}_t\pi_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0.99 & -0.967 & 0 \\ 0 & 0.99 & 0 & 0 \\ 0 & 0 & 0.99 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} rn_{t-1} \\ r_{t-1} \\ \pi_{t-1} \\ \mathbb{E}_{t-1}\pi_t \end{bmatrix} + \begin{bmatrix} 1.99 \\ 1 \\ 1 \\ 0.99 \end{bmatrix} \varepsilon_t.$$

The computation of the non-trivial transpose kernel of $A^\top$ yields

$$M = v^\top = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -0.633 & 0.633 & -0.445 & 0 \end{bmatrix} \in ker^\top\left(A^\top\right).$$

It follows that

$$Mx_t = y_t \longleftrightarrow$$
$$\longleftrightarrow \begin{bmatrix} 0 & 0 & 0 & 1 \\ -0.633 & 0.633 & -0.445 & 0 \end{bmatrix} \begin{bmatrix} rn_t \\ r_t \\ \pi_t \\ \mathbb{E}_t\pi_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbb{E}_t\pi_{t+1} \\ -0.633rn_t + 0.633r_t - 0.445\pi_t \end{bmatrix}.$$

Consequently, the minimal LTI state space representation guaranteeing a non-zero $VMA(0)$ of x_t in y_t is

$y_t = MB\varepsilon_t = D\varepsilon_t \longleftrightarrow$

$$\longleftrightarrow \begin{bmatrix} \mathbb{E}_t\pi_{t+1} \\ -0.633rn_t + 0.633r_t - 0.445\pi_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -0.633 & 0.633 & -0.445 & 0 \end{bmatrix} \begin{bmatrix} 1.99 \\ 1 \\ 1 \\ 0.99 \end{bmatrix} \varepsilon_t = \begin{bmatrix} 0.99 \\ -1.072 \end{bmatrix} \varepsilon_t,$$

whereby the four non-minimal states rn_t, r_t, π_t and $\mathbb{E}_t\pi_{t+1}$ are empirically represented by the two linear combinations (i.e. in principle and effective) $m_{11}rn_t + m_{12}r_t + m_{13}\pi_t + m_{14}\mathbb{E}_t\pi_{t+1} = \mathbb{E}_t\pi_{t+1}$ and $m_{21}rn_t + m_{22}r_t + m_{23}\pi_t + m_{24}\mathbb{E}_t\pi_{t+1} = -0.633rn_t + 0.633r_t - 0.445\pi_t$, the first three being actually measurable, the fourth being potentially so and to be all themselves represented as pseudo-variance shifted white noise $0.99\varepsilon_t$ and $-1.072\varepsilon_t$, respectively.

5.4 DSGE and LRE models. Even though it be not the sole empirical representation non-minimal states admit of, that obtained through the exploitation of the non-trivial transpose kernel of $A^\top$, $v^\top \neq 0$, is categorical and the probabilistically surest and is as such particularly problematic for first order DSGE and LRE models more in general, causing them to depart even further from reality at large (i.e. metaphysics).

To be sure, LRE models, to which DSGE belong, are normally ones of optimising representative agents, whose first order approximations express them as transition equations of (minimal) LTI state space representations in discrete and continuous time.

DSGE and LRE models in discrete time thus represented, as non-zero $VMA\,(0)$s of x_t in y_t, that is to say, can be consequently said to be convolutions of observables at best, however perfectly may such observables be measured; nay, they can be said to be no more than pseudo-variance shifted vectors of white noises for which no kind of forecast analysis is at all conceivable. It follows that DSGE and LRE models more in general, as normally conceived, can be viewed as consistently expedient as of second order approximations alone.

5.5 Empirical and structural models. Empirical VARs probably reflect an underlying structural model, additionally, however elementary it be, on account of the likely satisfaction of some mPMIC. Regardless of the sensation whereby, rather than forwards by trial and error, it may be best elaborated backwards, following the empirical VAR's parametric estimation and subsequent determination of A, B, C and D, duly restricted by theory, it would thereby appeal to the set of its probabilistically negligible empirical representations.

Such is nonetheless intuitable, for even though empirical models may ontologically precede structural models the latter logically precede the former, that is, they broaden their relations from forecasting to policy and thence seek validation through their replication (i.e. scientific method), so that the opposite should not occur.

Sims [15] had contended that the absence of a VAR representation of a structural model's states in the same's outputs need not invalidate the use of VARs to the end of recovering the inputs proper to the underlying structural model.

One hereby extends his contention by asserting that not only can empirical VARs be used in spite of such an absence but that the underlying structural models to which empirical VARs refer are almost never ones arising from transition equations, whose probabilistically surest empirical representations are by contrast considerably sterile.

Empirical VARs therefore result as being more comprehensively useful than suchlike structural models after all, that is, they result as presenting greater signification and expedience to the end of both forecasting and policy, despite the latter's phenomenal acceptation, for what can the soundness be that of policy indications elicited from structural models whose connexion with empirics is severely, if not wholly, sterile, whereby inputs are the shifted outputs themselves?

In order for such policy indications to be sound suchlike structural models can only be salvaged by appealing to axiomatic abstraction, although thereby conceding the essentialistic point held by authentic meta-physicists all along (i.e. not by subjectivistic neo-Keynesians and neo-monetarists behaving as phenomenalists and attacking reason under the banner of essentialism, being the Lucas critique).

5.6 Weaknesses of common LRE models. DSGE and LRE models' initial departure from reality at large, albeit conceptually self-inflicted, is in fact rooted in their potential failure to preserve their micro-foundations, on account of their subjection to the "Anything goes"[2] theorem whereby the canonical laws of supply and demand need not arise from their conceptual process of aggregation, speaking to Lippi's [12] contribution as well.

The linearisation of DSGE and LRE models to the end of deriving their solutions additionally causes them to lose their non-linear idiosyncrasies. A coherent alternative would therefore involve the adoption of macro-foundations as well as the numerical computation of their solutions in relation to their non-linear form.

While discretionally preserving optimisation, macro-foundations would permit one to conceptually dispose of representative agents and thereby enable one to meaningfully conduct parametric estimation by means of macroeconomic data, as per Kirman's [10] observation, notwithstanding the objections underlying their construction (e.g. imprecision, assumption violation). Accordingly, numerical solutions of their non-linear form would conceptually befit popular Bayesian parametrisation to a greater degree because of the preservation of all non-linear idiosyncrasies.

5.7 Companion matrices of first order LRE model unique solutions. In fact, observe that first order LRE model unique solutions present stable companion matrices by construction (see Examples 5.2 and 5.3): *ceteris paribus,* $\forall A$ in $x_t = Ax_{t-1} + Bu_t$ stemming from first order LRE model unique solutions, $|\lambda_{A(\lambda)}| < 1$, by construction.

Now, a matrix is invertible if and only if its determinant is non-zero and because the determinant of a matrix is the product of its eigenvalues according to their multiplicity a non-zero determinant signifies the absence of zero eigenvalues for a given matrix.

Since the invertibility of the companion matrices of first order LRE model unique solutions would imply a zero representation of the transition equation's non-minimal states a crucial question is thus whether their companion matrices may categorically present non-zero stable eigenvalues.

Otherwise stated, whenever the companion matrices of first order LRE model unique solutions be invertible there exists a zero representation of the transition equation's non-minimal states, which is additionally almost surely verified: *ceteris paribus,* $\lambda_{A(\lambda)} \neq 0 \longleftrightarrow det\,(A) \neq 0 \longleftrightarrow \exists A^{-1} \longleftrightarrow \exists \left(A^{\top}\right)^{-1} \longrightarrow 0 = v^{\top} = M = Mx_t = y_t$ for $A^{\top}v = 0$, almost surely, in which $\lambda_{A(\lambda)} \neq 0$ need not be.

Insofar as they were expressed in terms of LTI state space representations, suchlike first order LRE model unique solutions would not empirically exist: they would not phenomenally exist, that is, they would fail to exist both observationally and observably.

In other words, because of the stability properties of their companion matrices, the probabilistically surest empirical representation of suchlike first order LRE model unique solutions would be none other than zero, whereby they would not empirically exist.

The prescription would thus be that suchlike first order LRE models be studied by contemplating instability (unsteady state dynamics, long run demand effects etc.) and be therefore expressed as transition equations otherwise (i.e. by some other algorithm) or, afresh, that suchlike LRE models be studied as of their second order approximations at the minimum, if not in their non-linear forms outright.

For completeness, common algorithms for first order LRE model unique solutions, to which the question regarding the invertibility of their companion matrices applies, are those of Blanchard and Kahn [3], Anderson and Moore [1], Binder and Pesaran [2], King and Watson [9], Uhlig [16], Klein [11], Sims [14] and Christiano [4].

For instance, the companion matrices of first order LRE model unique solutions delivered by the algorithms of Blanchard and Kahn [3] and Sims [14] are $A = \left[(A_{11}\;0)\;\;(A_{21}\;0)\right]^{\top}$ and $A = \left[(A_{11}\;0)\;\;(0\;0)\right]^{\top}$, respectively.

Since both their determinants are zero (i.e. $det\,(A) = A_{11}\,(0) - (0)\,A_{21} = 0$ and $det\,(A) = A_{11}\,(0) - 0^2 = 0$) the companion matrices are non-invertible and the probabilistically surest empirical representation of their first order LRE model unique solutions is not zero, but a $VMA\,(0)$ of x_t in y_t.

6. Conclusion

[2] https://en.wikipedia.org

This work presented a systematic manner to derive a finite VAR representation for any minimal transition equation. It nevertheless clarified that such a representation, on account of the procedure underlying its derivation, is both the probabilistically surest and empirically useless, ranging from linear combinations of non-minimal states in principle, equal to shifted white noises, to output nullity, thereby presenting negative repercussions with particular regard to first order LRE models of optimising representative agents.

LRE models thus rendered could be consequently said to be convolutions of observables at best, however perfectly may such observables be measured. In brief, the probabilistically surest empirical representation of first order (approximations of) LRE models and of DSGE ones in particular is not merely a linear combination of non-minimal states in principle, whenever not equal to zero, but no more than a pseudo-variance shifted vector of white noises for which no kind of forecast analysis is at all conceivable.

References

[1] Anderson G. and Moore G. (1985) "A linear algebraic procedure for solving linear perfect foresight models", Economics letters.

[2] Binder M. and Pesaran M. (1994) "GAUSS and Matlab codes for multivariate rational expectations models and macroeconometric modelling: a review and some new results," Quantitative macroeconomics and real business cycles (codes 74).

[3] Blanchard O. and Kahn C. (1980) "The solution of rational linear difference models under rational expectations", Econometrica.

[4] Christiano L. (2022) "Solving dynamic equilibrium models by a method of undetermined coefficients", Computational economics.

[5] Fernández-Villaverde J., Rubio-Ramírez J., Sargent T. and Watson M. (2007) "ABCs (and Ds) of understanding VARs", American economic review.

[6] Franchi M. (2013) "Comment on: Ravenna F. 2007. Vector autoregressions and reduced form representations of DSGE models. Journal of Monetary Economics 54, 2048-2064.", Dipartimento di scienze statistiche empirical economics and econometrics working papers series.

[7] Franchi and Paruolo P. (2014) "Minimality of state space solutions of DSGE models and existence conditions for their VAR representation", Computational economics.

[8] Franchi and Vidotto A. (2013) "A check for finite order VAR representations of DSGE models", Economics letters.

[9] King R. and Watson M. (1998) "The solution of singular linear difference systems under rational expectations", International economic review.

[10] Kirman A. (1992) "Whom or what does the representative individual represent?", Journal of economic perspectives.

[11] Klein P. (2000) "Using the generalized Schur form to solve a multivariate linear rational expectations model", Journal of economic dynamics and control.

[12] Lippi M. (1988) "On the dynamic shape of aggregated error correction models", Journal of economic dynamics and control.

[13] Ravenna F. (2007) "Vector autoregressions and reduced form representations of DSGE models", Journal of monetary economics.

[14] Sims C. (2001) "Solving linear rational expectations models", Computational economics.

[15] Sims E. (2012) "News, non-invertibility, and structural VARs", Advances in econometrics.

[16] Uhlig H. (1999) "A toolkit for analyzing nonlinear dynamic stochastic models easily", in Marimon R. and Scott A., Computational methods for the study of dynamic economies.

Appendix

Julia commands for Example 5.2 (wherein # must replace % and α and β must replace alpha and beta)

```
using LinearAlgebra

% Parameters
```

```
alpha=0.6; % Expected inflation coefficient
beta=0.4; % Past inflation coefficient
gamma=0.01; % Exogenous shock coefficient

% LRE model solution: Q*x_t=R*x_t-1+S*varepsilon_t+T*eta_t
Q=[1 -alpha; 1 0]; % x_t=[pi_t; E_t(pi_t+1)]
R=[beta 0; 0 1];

S=[gamma; 0];
T=[0; 1]; % eta_t=pi_t-E_t-1(pi_t)

% LRE model solution unicity and stability
F=schur(Q, R); % Generalised Schur decomposition: H*J_Q*K'x_t=H*J_R*K'*x_t-1+S* ...
    varepsilon_t+T*eta_t, where H=F.Q and K=F.Z and upper triangular J_Q=F.S and J_R=F.T ...
    such that F.Q*F.S*F.Z'=Q and F.Q*F.T*F.Z'=R
select=abs.(F.beta./F.alpha).<1; % Logical vector for generalised eigenvalues reordering ...
    in ascending order, where F.alpha are J_Q'eigenvalues and F.beta are J_R's eigenvalues
% select=abs.(diag(F.T)./diag(F.S)).<1; % Logical vector for generalised eigenvalues ...
    reordering in ascending order, where J_Q=F.S and J_R=F.T
G=ordschur(F, select); % Schur decomposition with generalised eigenvalues reordered in ...
    ascending order, where J_Q=G.S and J_R=G.T
if (sum(abs.(G.beta./G.alpha).<1)==1) % Solution unicity check: n(generalised stable ...
    eigenvalues)=n(past variables)
% if (sum(abs.(diag(G.T)./diag(G.S)).≥1)==1) % Solution unicity check: n(generalised ...
    unstable eigenvalues)=n(expectational errors, future variables)
    println("There exists a unique and stable model solution.");
else
    println("There does not exist a unique and stable model solution.");
end

% State equation: ...
    x_1t=(JQ11^-1*JR11)*x_t-1+(K11*JQ11^-1)*(U1-V1*V2^-1*U2)*varepsilon_t=A11*x_t-1+B1*varepsilon_t ...
    and x_2t=(K21*JQ11^-1)*(U1-V1*V2^-1*U2)*varepsilon_t=B2*varepsilon_t, where ...
    U1=Ht11*S1+Ht12*S2, U2=Ht21*S1+Ht22*S2, V1=Ht11*T1+Ht12*T2, V2=Ht21*T1+Ht22*T2
JQ11=G.S[1:1, 1:1];
JR11=G.T[1:1, 1:1];
K11=G.Z[1:1, 1:1];
K21=G.Z[2:2, 1:1];
Ht11=G.Q'[1:1, 1:1];
Ht12=G.Q'[1:1, 2:2];
Ht21=G.Q'[2:2, 1:1];
Ht22=G.Q'[2:2, 2:2];

S1=S[1:1, 1:1];
S2=S[2:2, 1:1];
T1=T[1:1, 1:1];
T2=T[2:2, 1:1];

U1=Ht11*S1+Ht12*S2;
U2=Ht21*S1+Ht22*S2;
V1=Ht11*T1+Ht12*T2;
V2=Ht21*T1+Ht22*T2;

A11=inv(JQ11)*JR11;
B1=K11*inv(JQ11)*(U1-V1*inv(V2)*U2);
B2=K21*inv(JQ11)*(U1-V1*inv(V2)*U2);
A=[A11 0; 0 0]; % Companion matrix
B=[B1; B2]; % Shock matrix

% State equation stability
evalA, evecA=eigen(A);
if (all(abs.(evalA).<1)) % Stability check: A modulus eigenvalues<1
    println("The system is stable.");
else
    println("The system is unstable.");
end
```

```
% Transpose kernel of A^T for C=0
kerAt=nullspace(transpose(A)); % Kernel of A^T
Ms=transpose(kerAt); % Transpose kernel of A^T
Ds=Ms*B % Shock matrix of observables for C=0

transpose(Ms*A)==transpose(A)*kerAt % (v^T*A)^T=A^T*v
```

JULIA commands for Example 5.3 (wherein # must replace % and α and β must replace alpha and beta)

```
using LinearAlgebra

% LRE model solution: Q*x_t=R*x_t-1+S*varepsilon_t+T*eta_t
Q=[1 -1 0 -1; 0 1 0 0; 0 0 1 0; 0 0 1 0]; % x_t=[rn_t; r_t; pi_t; E_t(pi_t+1)]
R=[0 0 0 0; 0 0.99 0 0; 0 0 0.99 0; 0 0 0 1];

S=[0; 1; 1; 0];
T=[0; 0; 0; 1]; % eta_t=pi_t-E_t-1(pi_t)

% LRE model solution unicity and stability
F=schur(Q, R); % Generalised Schur decomposition: H*J_Q*K'x_t=H*J_R*K'*x_t-1+S* ...
    varepsilon_t+T*eta_t, where H=F.Q and K=F.Z and upper triangular J_Q=F.S and J_R=F.T ...
    such that F.Q*F.S*F.Z'=Q and F.Q*F.T*F.Z'=R
select=abs.(F.beta./F.alpha).<1; % Logical vector for generalised eigenvalues reordering ...
    in ascending order, where F.alpha are J_Q'eigenvalues and F.beta are J_R's eigenvalues
% select=abs.(diag(F.T)./diag(F.S)).<1; % Logical vector for generalised eigenvalues ...
    reordering in ascending order, where J_Q=F.S and J_R=F.T
G=ordschur(F, select); % Schur decomposition with generalised eigenvalues reordered in ...
    ascending order, where J_Q=G.S and J_R=G.T
if (sum(abs.(G.beta./G.alpha).<1)==3) % Solution unicity check: n(generalised stable ...
    eigenvalues)=n(past variables)
% if (sum(abs.(diag(G.T)./diag(G.S)).≥1)==1) % Solution unicity check: n(generalised ...
    unstable eigenvalues)=n(expectational errors, future variables)
    println("There exists a unique and stable model solution.");
else
    println("There does not exist a unique and stable model solution.");
end

% State equation: ...
    x_1t=(JQ11^-1*JR11)*x_t-1+(K11*JQ11^-1)*(U1-V1*V2^-1*U2)*varepsilon_t=A11*x_t-1+B1*varepsilon_t ...
    and x_2t=(K21*JQ11^-1)*(U1-V1*V2^-1*U2)*varepsilon_t=B2*varepsilon_t, where ...
    U1=Ht11*S1+Ht12*S2, U2=Ht21*S1+Ht22*S2, V1=Ht11*T1+Ht12*T2, V2=Ht21*T1+Ht22*T2
JQ11=G.S[1:3, 1:3];
JR11=G.T[1:3, 1:3];
K11=G.Z[1:3, 1:3];
K21=G.Z[4:4, 1:3];
Ht11=G.Q'[1:3, 1:3];
Ht12=G.Q'[1:3, 4:4];
Ht21=G.Q'[4:4, 1:3];
Ht22=G.Q'[4:4, 4:4];

S1=S[1:3, 1:1];
S2=S[4:4, 1:1];
T1=T[1:3, 1:1];
T2=T[4:4, 1:1];

U1=Ht11*S1+Ht12*S2;
U2=Ht21*S1+Ht22*S2;
V1=Ht11*T1+Ht12*T2;
V2=Ht21*T1+Ht22*T2;

A11=inv(JQ11)*JR11;
B1=K11*inv(JQ11)*(U1-V1*inv(V2)*U2);
B2=K21*inv(JQ11)*(U1-V1*inv(V2)*U2);
A=[A11 zeros(3, 1); zeros(1, 3) 0]; % Companion matrix
B=[B1; B2]; % Shock matrix
```

```
% State equation stability
evalA, evecA=eigen(A);
if (all(abs.(evalA).<1)) % Stability check: A modulus eigenvalues<1
    println("The system is stable.");
else
    println("The system is unstable.");
end

% Transpose kernel of A^T for C=0
kerAt=nullspace(transpose(A)); % Kernel of A^T
Ms=transpose(kerAt); % Transpose kernel of A^T
Ds=Ms*B % Shock matrix of observables for C=0

transpose(Ms*A)==transpose(A)*kerAt % (v^T*A)^T=A^T*v
```