14.384 Time Series Analysis, Fall 2007
Professor Anna Mikusheva
Paul Schrimpf, scribe
September 25, 2007
Corrected October, 2012 Lecture 10

Introduction to VARs

Wold Decomposition Theorem

Theorem 1 (Wold decomposition). *Let y_t be a 2nd order stationary process with $Ey_t = 0$. Then y_t could be written as $y_t = v_t + c(L)e_t$, where*

1. *e_t: $Ee_t = 0$, $Ee_te_t' = \Omega$, and $e_t \in span\{y_t, y_{t-1}, ...\} = I_t$ (i.e. e_t are fundamental)*
2. *$Ee_te_s' = 0$ for all $t \neq s$*
3. *$\sum_{j=0}^{\infty} \|c_j\|^2 < \infty$*
4. *v_t is deterministic, that is $v_t \in I_{-\infty} = \cup_{j=0}^{\infty} I_{t-j}$*

Remark 2. The span$\{y_t, y_{t-1}, ...\} = I_t$ is the space of all linear combinations of $\{y_t, y_{t-1}, ...\}$.

Proof. The theorem is stated for vector processes, but we'll just prove it in the scalar case.
We will prove the theorem by constructing e_t and showing that it satisfies the conditions stated. Let $e_t = y_t - \hat{E}(y_t|I_{t-1}) = y_t - a(L)y_{t-1}$, where $\hat{E}(y_t|I_{t-1})$ is the best linear forecast of y_t from past y. Note that the form of this forecast ($a(L)$) does not depend on t. This is because y_t is second order stationary, and the best linear forecast only depends on the first and second moments of y_t. By best linear forecast, we mean that $a(L)$ solves:

$$\min_{\{a_j\}} E(y_t - \sum_{j=1}^{\infty} a_j y_{t-j})^2$$

The first order conditions are

$$\begin{aligned} \frac{\partial}{\partial a_j} : \; E[y_{t-j}(y_t - \sum_{j=1}^{\infty} a_j y_{t-j})] = 0 \\ E[y_{t-j}e_t] = 0 \end{aligned}$$

So e_t is uncorrelated with y_{t-j} $\forall j > 0$. This implies that $I_t = I_{t-1} \oplus \{e_t\}$.
We now verifty that e_t satisfies conditions 1-4.

1. $Ey_t = 0$ $\forall t$ implies that $Ee_t = 0$.
   Also, $\text{Var}(e_t) = E(y_t - a(L)y_t)^2$ is just a function of the covariances of y_t and a_j, so $\text{Var}(e_t) = \sigma^2$ is constant.
2. $Ee_ty_{t-j} = 0$ $\forall j \geq 1$, and $e_{t-j} = y_{t-j} - a(L)y_{t-j}$ implies $Ee_te_{t-j} = 0$.
3. By definition of e_t,

$$y_t = e_t + \sum_{k=1}^{\infty} a_k y_{t-k}$$

Repeatedly substituting for y_{t-k} gives

$$\begin{aligned}
y_t =& e_t + \sum_{k=1}^{\infty} a_k y_{t-k} \\
=& e_t + a_1(e_{t-1} + \sum_{k=1}^{\infty} a_k y_{t-1-k}) + \sum_{k=2}^{\infty} a_k y_{t-k} \\
=& e_t + a_1 e_{t-1} + \sum_{k=1}^{\infty} \tilde{a}_k y_{t-k-1} \\
=& e_t + a_1 e_{t-1} + v_t^1 \\
\vdots& \\
=& \sum_{j=0}^{k} c_j e_{t-j} + v_t^k
\end{aligned}$$

where $v_t^k \in \text{span}\{y_{t-k-1}, y_{t-k-2}, ...\}$. Consider:

$$\begin{aligned}
E(y_t - \sum_{j=0}^{k} c_j e_{t-j})^2 =& E y_t^2 + \sum_{j=0}^{k} c_j^2 \sigma^2 - 2 \sum_{j=0}^{k} c_j E[y_t e_{t-j}] \\
=& E y_t^2 + \sum_{j=0}^{k} c_j^2 \sigma^2 - 2 \sum_{j=0}^{k} c_j^2 \sigma^2 \\
=& E y_t^2 - \sum_{j=0}^{k} c_j^2 \sigma^2
\end{aligned}$$

where we used the fact that $E[y_t e_{t-j}] = E[(\sum_{l=0}^{k} c_l e_{t-l} + v_t^k) e_{t-j}] = c_l \sigma^2$. Now we know that $E(y_t - \sum_{j=0}^{k} c_j e_{t-j})^2 \geq 0$. Therefore, it must be that $\sum_{j=0}^{k} c_j^2 \leq \frac{E y_t^2}{\sigma^2}$ and $\sum_{j=0}^{\infty} c_j^2 \leq \infty$. This implies that, $\sum_{j=0}^{\infty} c_j e_{t-j}$ is well-defined.

4. Finally, $v_t = y_t - \sum_{j=0}^{\infty} c_j e_{t-j} \in I_{-\infty}$.

$\square$

Discussion

Definition 3. If $v_t = 0$, the y_t is called *linearly regular.*

- Wold decomposition (an MA representation satisfying conditions (1)-(4) of the theorem)is "unique." We put unique in quotes, because it is only unique up to a rotation of e_t. Let R be a rotation. $e_t = R u_t$, then $y_t = c(L) R u_t = \tilde{c}(L) u_t$, where $\tilde{c}_j = c_j R$. What is really unique is the space spanned by $\{e_t\}$. Structural VARs are about how we pick a particular rotation of the, and then attach interpretations to the components of e_t.

- It's true that any 2nd order stationary process has an MA representation, but that doesn't mean that it's always a good idea to study the MA representation.

  Example 4. Two-state Markov chain: $\begin{cases} P(y_t = 1 | y_{t-1} = 1) = p \\ P(y_t = 0 | y_{t-1} = 0) = q \end{cases}$. This is 2nd order stationary and has the following MA representation:

$$
\begin{aligned}
E(y_t|y_{t-1}) =& \begin{cases} p & y_{t-1} = 1 \\ 1-q & y_{t-1} = 0 \end{cases} \\
=& 1 - q + y_{t-1}(q - 1 + p)
\end{aligned}
$$

so

$$
\begin{aligned}
y_t =& 1 - q + y_{t-1}(q - 1 + p) + e_t \\
=& \frac{1-q}{2-p-q} + \sum_{j=0}^{\infty}(p+q-1)^j e_{t-j}
\end{aligned}
$$

where

$$
e_t = \begin{cases} 1-p & \text{prob } p \,|\, y_{t-1} = 1 \\ -p & \text{prob } 1-p \,|\, y_{t-1} = 1 \\ -(1-q) & \text{prob } q \,|\, y_{t-1} = 0 \\ q & \text{prob } 1-q \,|\, y_{t-1} = 0 \end{cases}
$$

The MA representation is not the best description of the dynamics in this case. This example is simple, but the point is more general. There are macro models with simple state-space representations and complicated MA representations.

- There are many MA representations. The Wold representation is only unique up to rotation when you require the MA to be fundamental.

  Example 5. Suppose $y_t = e_t + \theta e_{t-1}$, $\text{Var}(e_t) = \sigma^2$. Then $\gamma_0 = (1+\theta^2)\sigma^2$, $\gamma_1 = \theta\sigma^2$, $\gamma_j = 0 \ \forall j > 1$. Consider $y_t = \eta_t + \frac{1}{\theta}\eta_{t-1}$ with $\text{Var}\eta_t = \theta^2\sigma^2$. It is easy to see that this process has the same covariances. However, only the first representation is fundamental.

$$
\begin{aligned}
y_t =& (1+\theta L)e_t \\
e_t =& (1+\theta L)^{-1} y_t = \sum_{j=0}^{\infty}(-\theta)^j y_{t-j}
\end{aligned}
$$

  For the second process, first define $F = L^{-1}$.

$$
\begin{aligned}
y_t =& (\theta F + 1)\eta_{t-1} \\
\eta_{t-1} =& \theta \sum_{j=0}^{\infty}(-\theta)^j F^j y_t
\end{aligned}
$$

  Remark 6. It is possible that your model has a non-fundamental MA representation. A fundamental representation requires that your shocks are spanned by the set of variables in your VAR. For example, if you're looking at shocks to money supply in a VAR with money and output, it is very likely that the Fed looks at variables other than output and money and when choosing policy, so shocks to the money supply may not be fundamental. One solution might be to include more variables in your VAR. Relatedly, you might use a factor-augmented VAR.

Remark 7. **Relation to spectrum decomposition**: there are three ways to represent a process: the Wold decomposition, the covariances, and the spectrum. There is a 1-1 mapping among these representations, in the following sense: there is a one-to-one mapping between the spectrum and the set of auto-covariances; there is also a 1-1 mapping between autocovariances and the set of coefficients of the fundamental MA representation.