Why Are Vector Autoregressions Useful in Finance?

University of Rome at Tor Vergata

Fall 2001

Gerald P. Dwyer, Jr.

Any views are the author’s and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

Overview of Vector Autoregressions

- Vector autoregressions are
  - A concise way of summarizing data
  - Generally have little serial correlation in residuals
  - Can be used to examine complex relationships among variables
- Examples
  - Information content in prices in different asset markets
  - Effects of arbitrage across markets
  - Relationships between options and cash markets

Truth in Advertising

- What a Vector Autoregression (VAR) cannot do

  - It generally is not an intelligent causal model of the data

    - A VAR can summarize the data

    - Generally is not the set of equations implied by any theory

Vector Autoregression Illustrated

- A simple vector autoregression for two variables *y* and *z* is

$$y_t = \alpha^y + \beta^y y_{t-1} + \gamma^y z_{t-1} + \varepsilon_t^y$$
$$z_t = \alpha^z + \beta^z y_{t-1} + \gamma^z z_{t-1} + \varepsilon_t^z$$

  - where

    - the α's and β's are parameters

    - the epsilons are white noise, i.,e.,

    -

$$\mathrm{E}\,\varepsilon_t^i = 0, \mathrm{Var}\,\varepsilon_t^i = \sigma^2 \text{ and } \mathrm{Cov}(\varepsilon_t^i, \varepsilon_s^j) = 0$$
$$i, j = y, z \text{ and } i \neq j, t \neq s$$

Vector Autoregression
Cleaner Representation

- Matrix notation

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \alpha^y \\ \alpha^z \end{bmatrix} + \begin{bmatrix} \beta^y & \gamma^y \\ \beta^z & \gamma^z \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}$$

  - which can be written

$$\mathbf{x_t} = \alpha + \beta \mathbf{x_{t-1}} + \varepsilon_{\mathbf{t}}$$

  - which is a vector autoregression
    - uses vectors
    - is an autoregression in the vectors

Vector Autoregression
Generalizes Easily

- More variables
  - Increase the number of rows and columns
  - Illustration
    - $\mathbf{x_t} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{x_{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}$
      - Vectors (2 x 1)
      - Matrix $\boldsymbol{\beta}$ (2 x 2)
  - Could just as well be
    - $\mathbf{x_t}$, $\mathbf{x}_{\mathbf{t}-1}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\varepsilon}_\mathbf{t}$
      - (3 x 1), (4 x 1), ... (n x 1)
    - $\boldsymbol{\beta}$
      - (3 x 3), (4 x 4), ... (n x n)

Vector Autoregression

Generalizes Easily

- More lags of variables

$$y_t = \alpha^y + \beta_1^y y_{t-1} + \beta_2^y y_{t-2} + \gamma_1^y z_{t-1} + \gamma_2^y z_{t-2} + \varepsilon_t^y$$

$$z_t = \alpha^z + \beta_1^z y_{t-1} + \beta_2^z y_{t-2} + \gamma_1^z z_{t-1} + \gamma_2^z z_{t-2} + \varepsilon_t^y$$

- Two ways
  - Polynomials in lag operator
  - Augmented matrices

Augmented Matrices

- Original equations

$$y_t = \alpha^y + \beta_1^y y_{t-1} + \beta_2^y y_{t-2}$$
$$+ \gamma_1^y z_{t-1} + \gamma_2^y z_{t-2} + \varepsilon_t^y$$
$$z_t = \alpha^z + \beta_1^z y_{t-1} + \beta_2^z y_{t-2}$$
$$+ \gamma_1^z z_{t-1} + \gamma_2^z z_{t-2} + \varepsilon_t^y$$

  - Want

$$\mathbf{x_t} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{x_{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}$$

  - Add variables in $\mathbf{x_t}$ so that have $\mathbf{x_t}$ and $\mathbf{x_{t-1}}$ progress naturally in time
    - $\mathbf{x_{t-1}}$ will become $\mathbf{x_t}$ next period

Augmented Matrices Construction

- Add variables to $\mathbf{x_t}$ and $\mathbf{x_{t-1}}$

  Augment $\mathbf{x_t}$ and $\mathbf{x_{t-1}}$

  - $\mathbf{x_{t-1}}$ has to include $y_{t-1}, y_{t-2}, z_{t-1}, z_{t-2}$

    - So include them

    - $\mathbf{x_{t-1}}' = [y_{t-1}, y_{t-2}, z_{t-1}, z_{t-2}]$

  - Implies that

    - $\mathbf{x_t}' = [y_t, y_{t-1}, z_t, z_{t-1}]$

  - What is vector $\boldsymbol{\alpha}$?
  - What is matrix $\boldsymbol{\beta}$?

Contents of Augmented Matrices

- Just fill in the blanks in **α** and **β**

$$\begin{bmatrix} y_t \\ y_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} \alpha^y \\ 0 \\ \alpha^z \\ 0 \end{bmatrix}$$

$$+\begin{bmatrix} \beta_1^y & \beta_2^y & \gamma_1^y & \gamma_2^y \\ ? & ? & ? & ? \\ \beta_1^z & \beta_2^z & \gamma_1^z & \gamma_1^z \\ ? & ? & ? & ? \end{bmatrix}\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ z_{t-1} \\ z_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ 0 \\ \varepsilon_t^z \\ 0 \end{bmatrix}$$

Contents of Augmented Matrices Completed

$$\begin{bmatrix} y_t \\ y_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} \alpha^y \\ 0 \\ \alpha^z \\ 0 \end{bmatrix}$$

$$+ \begin{bmatrix} \beta_1^y & \beta_2^y & \gamma_1^y & \gamma_2^y \\ - & - & - & - \\ \beta_1^z & \beta_2^z & \gamma_1^z & \gamma_1^z \\ - & - & - & - \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ z_{t-1} \\ z_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ 0 \\ \varepsilon_t^z \\ 0 \end{bmatrix}$$

- Could get rid of $\boldsymbol{\alpha}$ by just including constant in variables in $\mathbf{x}$
  - We'll just suppress constants
  - Deviations from means would let us ignore constants

Properties of First-order Representation

- Matrix $\boldsymbol{\beta}$ is not singular
  - despite the augmentation

- If the ε's are serially uncorrelated, then ordinary least squares is a consistent estimator of

  - the parameters in $\boldsymbol{\beta}$

- If the ε's are serially uncorrelated and homoskedastistic, then ordinary least squares is a consistent estimator of

  - the variance of the ε's

  - the serial correlation of the ε's

- Can say these things without worrying about "stationarity" and "unit roots"

Unit Roots

- What are the implications of unit roots?

- Unibiasedness
-
  - Proving unbiasedness not an issue

$$\mathrm{E}\,\beta^{ols}_{x_t, x_{t-1}} = \mathrm{E}\frac{\sum x_t x_{t-1}}{\sum x_{t-1}^2}$$

$$\neq \frac{\mathrm{E}\sum x_t x_{t-1}}{\mathrm{E}\sum x_{t-1}^2}$$

  - whereas

$$\operatorname{plim}\beta^{ols}_{x_t, x_{t-1}} = \operatorname{plim}\frac{\sum x_t x_{t-1}}{\sum x_{t-1}^2}$$

$$= \frac{\operatorname{plim}\sum x_t x_{t-1}}{\operatorname{plim}\sum x_{t-1}^2}$$

Unit Roots

- Suppose no unit roots
  - Additional assumptions
    - Homoskedastic errors
    - No serial correlation of errors
  - Ordinary least squares estimator of $\boldsymbol{\beta}$
  - Consistent
  - OLS estimates of standard errors are correct in the sense that
    - Estimated coefficients divided by their standard deviations are asymptotically normal under the null hypothesis that the coefficients are zero
    - Standard F-statistics have an F distribution

Unit Roots

- Suppose that variables have unit roots

  - e.g. $x_t = x_{t-1} + \varepsilon_t^x$

  - Then coefficient of at least one variable relative to its standard deviation will not have a normal distribution

- Solution depends on issue of cointegration

Cointegration

- Definition of cointegrated variables
  - Suppose that y and z have unit roots
  - y and z are cointegrated if, e.g., $y_t - \delta z_t$ does not have a unit root ($\delta \neq 0$)

Estimation Depending on Cointegration

- Suppose that variables **not cointegrated**
  - Estimate VAR in the first differences
  - Δx_t and Δy_t wherever *x* and *y* appear above

- Suppose that variables are **cointegrated**
  - Estimate Vector Error Correction Mechanism (VECM)
    - VAR with additional terms for cointegrating vectors

Vector Error Correction Mechanism

- VECM with two variables and two lags is

$$\Delta y_t = \lambda_y (y_{t-1} - \delta z_{t-1}) + \beta_1^y \Delta y_{t-1} + \beta_2^y \Delta y_{t-2}$$

$$+ \gamma_1^y \Delta z_{t-1} + \gamma_2^y \Delta z_{t-2} + \varepsilon_t^y$$

$$\Delta z_t = \lambda_y (y_{t-1} - \delta z_{t-1}) + \beta_1^z \Delta y_{t-1} + \beta_2^z \Delta y_{t-2}$$

$$+ \gamma_1^z \Delta z_{t-1} + \gamma_2^z \Delta z_{t-2} + \varepsilon_t^y$$

  - Same as VAR with a term added that is the cointegrating vector

  - Why one cointegrating vector?
    Why not two?
    - If N variables all have unit roots, then there can be at most N-1 cointegrating vectors
    - Hence, one cointegrating vector here

Summary

- Many other details

- Basic idea is that a VAR or VECM summarizes the correlations in the data

- VAR or VECM is especially useful when variables are serially correlated

Applications

- Can use VAR or VECM to obtain estimate of the response of one variable when another changes
  - This is a tricky issue .
  - Simplest if correlation of errors is zero across equations in VAR or VECM
- Can use VAR or VECM to examine whether one variable helps to predict the other
  - “Granger causality”
  - Can answer substantive questions such as whether price on one of two markets reflects information faster

Applications

- Dividends and stock prices

  - Let d be the logarithm of real dividends and p be the logarithm of real stock prices

  - Suppose that dividends and stock prices are cointegrated with a coefficient of one - $d - p$ does not have a unit root

  - VECM for d and p

$$\Delta d_t = \lambda_y (d_{t-1} - p_{t-1}) + \beta_1^y \Delta d_{t-1} + \beta_2^y \Delta d_{t-2}$$

$$+ \gamma_1^y \Delta p_{t-1} + \gamma_2^y \Delta p_{t-2} + \varepsilon_t^y$$

$$\Delta p_t = \lambda_y (d_{t-1} - p_{t-1}) + \beta_1^z \Delta d_{t-1} + \beta_2^z \Delta d_{t-2}$$

$$+ \gamma_1^z \Delta p_{t-1} + \gamma_2^z \Delta p_{t-2} + \varepsilon_t^y$$

Applications (continued)

- Price of cash and future
  - For example, stock price index and futures price
    - Want high-frequency data for this
    - Probably are cointegrated with a coefficient of one
    - Similar VECM to one for dividends and stock prices
    - When done, may want more elaborate model that allows for on a nonlinear relationship

Applications (continued)

- Stock price index and volatility of stock index from options market
  - No obvious reason why they should be cointegrated
  - VAR would be more appropriate if not cointegrated
  - It might pay to look at the relationship between the implied volatility more carefully
    - Context of option-pricing models
    - Nonlinear models

Conclusion

- Vector autoregressions and Vector error correction mechanisms can be very useful

- They can be useful summaries of the relationship between variables

- They are not a substitute for thought

References

There are many suitable references for the level of generality of this lecture. One such reference is

Greene, William H. 2000. *Econometric Analysis*. Fourth edition. New York: Prentice Hall, Chapter 17.

Another reference that has more detail and more English is

Kennedy, Peter. 1998. *A Guide to Econometrics*. Fourth edition. Cambridge: The MIT Press, Chapter 17.