Applied Economic Forecasting

8. VAR Models

1. Vector Autoregessive (VAR) Models
2. Impulse Response Functions (IRF)
3. Forecast Error Variance Decomposition (FEVD)

Section 1

Vector Autoregessive (VAR) Models

VAR Models

What is it?

- Vector Autoregression (VAR) is a multivariate forecasting algorithm that is used when two or more time series are assumed to influence each other.
- Each time series is modeled as a function of the past values of itself and past lags of the other variables.

How do VARs differ from other Autoregressive models?

Unlike the AR, ARMA, or ARIMA models, these models allow for potential bi-directional relationships.

- The predictors (**X**s) influence Y and are possibly affected by Y as well.
- All variables are therefore treated as “endogenous”

Introduction

Using two random variables, $(y_{1t},\ y_{2t})'$ as an example. Then, a bivariate model of order 1 takes the following form:

$$y_{1t} = a_{10} + b_{12}y_{2,t} + \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \varepsilon_{1t} \tag{1}$$

$$y_{2t} = a_{20} + b_{21}y_{1,t} + \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + \varepsilon_{2t} \tag{2}$$

1. b_{12} and b_{21} represent ***contemporaneous/simultaneous/instantaneous effects of:***
   - i. a unit change of y_{2t} on y_{1t}
   - ii. and a unit change of y_{1t} on y_{2t} , respectively.
2. ϕ_{12} and ϕ_{21} represent ***feedback effects of:***
   - i. a unit change of $y_{2,t-1}$ on y_{1t}
   - ii. and a unit change of $y_{1,t-1}$ on y_{2t} , respectively.

$$y_{1t} = a_{10} + b_{12}y_{2,t} + \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \varepsilon_{1,t} \quad (3)$$
$$y_{2t} = a_{20} + b_{21}y_{1,t} + \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + \varepsilon_{2,t} \quad (4)$$

- $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise disturbances with standard deviations σ_1 and σ_2.
- $\{\varepsilon_{1,t}\}$ and $\{\varepsilon_{2,t}\}$ are uncorrelated.
- System incorporates feedback as y_1 and y_2 are allowed to affect each other:
  - ϕ_{ii} coefficients captures the influence of the lag of each variable on itself, while the ϕ_{ij} coefficients captures the influence of the lag of variable y_j on y_i.

VAR in matrix form

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{5}$$

- This can be rewritten as

$$\boldsymbol{A} Y_t = \boldsymbol{A_0} + \boldsymbol{A_1} Y_{t-1} + \boldsymbol{\varepsilon}_t \tag{6}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, \quad Y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, \quad Y_{t-1} = \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix}, \quad \boldsymbol{A_0} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix},$$
$$\boldsymbol{A_1} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Note

The matrix form VAR (Eqn 5) and the rewritten form (Eqn 6) are referred to as a **Structural Vector Autoregressive model (SVAR)** if at least one of the off-diagonal elements of $\boldsymbol{A}$ is nonzero and $\boldsymbol{\varepsilon}_t$ has a diagonal covariance matrix, Σ_ε.

VAR Models

- Transform the SVAR to its reduced VAR form by premultiplying by $\boldsymbol{A}^{-1}$,

$$Y_t = \boldsymbol{B_0} + \boldsymbol{B_1} Y_{t-1} + \boldsymbol{\upsilon}_t \tag{7}$$

where

$$\begin{aligned} \boldsymbol{B_0} &= \boldsymbol{A}^{-1}\boldsymbol{A_0} \\ \boldsymbol{B_1} &= \boldsymbol{A}^{-1}\boldsymbol{A_1} \\ \boldsymbol{\upsilon}_t &= \boldsymbol{A}^{-1}\boldsymbol{\varepsilon}_t \end{aligned}$$

Note

Y_t written this way is referred to VAR(1) where $\boldsymbol{\upsilon}_t$ is a white noise process.

VAR Stability and Stationarity

- Suppose we have an initial value Y_0, then we can apply backward iteration such that

$$
\begin{aligned}
Y_t &= \boldsymbol{B_0} + \boldsymbol{B_1} Y_{t-1} + \boldsymbol{v}_t \\
&= \boldsymbol{B_0} + \boldsymbol{B_1}(\boldsymbol{B_0} + \boldsymbol{B_1} Y_{t-2} + \boldsymbol{v}_{t-1}) + \boldsymbol{v}_t \\
&= \boldsymbol{B_0} + \boldsymbol{B_0}\boldsymbol{B_1} + \boldsymbol{B_1^2} Y_{t-2} + \boldsymbol{B_1}\boldsymbol{v}_{t-1} + \boldsymbol{v}_t \\
&\vdots \\
&= \left(I + \boldsymbol{B_1} + \ldots + \boldsymbol{B_1}^{t-1}\right)\boldsymbol{B_0} + \boldsymbol{B_1^t} Y_0 + \sum_{i=0}^{t-1} \boldsymbol{B_1^i} \boldsymbol{v}_{t-i}
\end{aligned}
$$

VAR Stability and Stationarity

Side Note I

- If $\{\lambda_1, \ldots, \lambda_k\}$ are the eigenvalues of $\boldsymbol{B_1}$, then $\{\lambda_1^t, \ldots, \lambda_k^t\}$ are the eigenvalues of $\boldsymbol{B_1^t}$.
- Also, if all eigenvalues of a matrix are 0, then the matrix must be $\emptyset$.
- So, if $\boldsymbol{B_1^t} \to 0$ as $t \to \infty$, then $\{\lambda_1^t, \ldots, \lambda_k^t\} \to 0$ as $t \to \infty$
- This implies that the absolute value of all eigenvalues of $\boldsymbol{B_1}$ must be less than 1.

VAR Stability and Stationarity

Side Note II

The eigenvalues of $\boldsymbol{B_1}$ are solutions of the determinant equation

$$|\lambda I - \boldsymbol{B_1}| = 0$$

which can be written as

$$|I - \boldsymbol{B_1}\frac{1}{\lambda}| = 0$$

$$|I - \boldsymbol{B_1}z| = 0$$

where $z = \frac{1}{\lambda}$ and should be greater than 1 in absolute value.

VAR Stability and Stationarity

So, as $t \to \infty$, if

$$det(I - \boldsymbol{B_1} z) = 0, \text{ for } |z| > 1,$$

then

$$\left(I + \boldsymbol{B_1} + \ldots + \boldsymbol{B_1}^{t-1}\right)_{|t \to \infty} \to (I - \boldsymbol{B_1})^{-1}$$
$$\boldsymbol{B_1^t}_{|t \to \infty} \to 0$$
$$\sum_{i=0}^{t-1} \boldsymbol{B_1^i} \boldsymbol{v}_{t-i} \text{ exists}$$

VAR Stability and Stationarity

- So,

$$\mathbf{E}(Y_t) = (I - \boldsymbol{B_1})^{-1}\boldsymbol{B_0}$$

and the autocovariance

$$\begin{aligned}
\Gamma_Y(h) &= \mathrm{Cov}(Y_t, Y_{t-h}) = \mathbf{E}(Y_t - \mu)(Y_{t-h} - \mu)' \\
&= \lim_{n\to\infty} \sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{B}^i \mathbf{E}(\boldsymbol{\upsilon}_{t-i}\boldsymbol{\upsilon}_{t-h-j}) \boldsymbol{B}^{j'} \\
&= \lim_{n\to\infty} \sum_{j=0}^{n} \boldsymbol{B}^{h+j} \Sigma_\upsilon \boldsymbol{B}^j \\
&= \sum_{i=0}^{\infty} \boldsymbol{B}^{h+j} \Sigma_\upsilon \boldsymbol{B}^j
\end{aligned}$$

because $\mathbf{E}(\boldsymbol{\upsilon}_t \boldsymbol{\upsilon}_s') = 0$ for $s \neq t$, $\mathbf{E}(\boldsymbol{\upsilon}_t \boldsymbol{\upsilon}_t') = \Sigma_\upsilon$ for all t, and $\boldsymbol{B}^\infty \to 0$.

VAR Stability and Stationarity

A VAR(1) process is stable if all eigenvalues of $\boldsymbol{B_1}$ have modulus less than 1, that is

$$det(I - \boldsymbol{B_1} z) = 0, \text{ for } |z| > 1$$

where $z = \frac{1}{\lambda}$.

VAR(p) Model

- In general, a VAR(p) model is expressed as

$$Y_t = \boldsymbol{B_0} + \boldsymbol{B_1} Y_{t-1} + \ldots + \boldsymbol{B_p} + \boldsymbol{\upsilon}_t$$

where

$$Y_t = \begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{Nt} \end{pmatrix}$$

and $\boldsymbol{\upsilon}_t$ is $N-$dimensional white noise or innovation process, ie,

$$\mathbf{E}(\boldsymbol{\upsilon}_t) = 0$$
$$\mathbf{E}(\boldsymbol{\upsilon}_t \boldsymbol{\upsilon}_t^{'}) = \Sigma_\upsilon$$
$$\mathbf{E}(\boldsymbol{\upsilon}_t \boldsymbol{\upsilon}_{t-s}^{'}) = 0, \text{ for } t \neq s$$

Intuition

Recall that a typical AR model is a linear combination of it's own lags. To that end, we use past values of the variable to predict current and future values.

A standard AR(p) model would look like:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

where ϕ_0 is the intercept and $\phi_j \; ; \; j \in \{1, p\}$ are the lagged coefficients of Y.

Intuition

In the case of the VAR model, we have several Y variables (for example the GDP of US and China) and each has its own equation. For simplicity, let us assume that we have a VAR(1) model and two GDP series (y_1 and y_2).

$$y_{1,t} = \gamma_1 + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \varepsilon_{1,t} \tag{8}$$

$$y_{2,t} = \gamma_2 + \phi_{21} y_{1,t-1} + \phi_{22,} y_{2,t-1} + \varepsilon_{1,t} \tag{9}$$

Each variable is a linear function of the lag 1 values for all variables in the set.

If we had three variables we could easily add an additional equation and also an extra regressor to each.

$$y_{1,t} = \gamma_1 + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \phi_{13} y_{3,t-1} + \varepsilon_{1,t} \tag{10}$$

$$y_{2,t} = \gamma_2 + \phi_{21} y_{1,t-1} + \phi_{22,} y_{2,t-1} + \phi_{23} y_{3,t-1} + \varepsilon_{1,t} \tag{11}$$

$$y_{3,t} = \gamma_2 + \phi_{21} y_{1,t-1} + \phi_{22,} y_{2,t-1} + \phi_{33} y_{3,t-1} + \varepsilon_{1,t} \tag{12}$$

Uses of VAR

The VAR model can be useful for

1. forecasting a collection of related variables where no explicit interpretation is required;
2. testing whether one variable is useful in forecasting another (the basis of Granger causality tests);
3. impulse response analysis, where the response of one variable to a sudden but temporary change in another variable is analysed;
4. forecast error variance decomposition, where the proportion of the forecast variance of each variable is attributed to the effects of the other variables.

VAR Estimation Steps

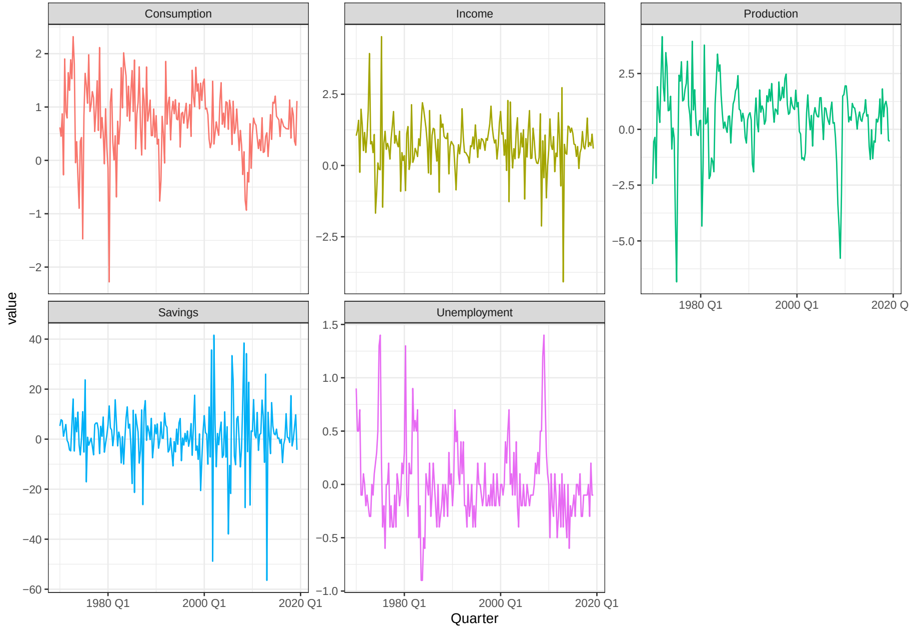
1. Visualize the Data.
2. Check for stationarity and make the time series stationary.

- Sims (1980) ans Sims, Stock, and Watson (1990) argue against differencing, even if the variables have a unit root.
    - Their argument: The role of VAR is to determine the relationship between the variables, not to determine the parameter values.
    - Similar argument against detrending.

3. Select the order (p) of the VAR Model
4. Estimate the implied model
5. Check for Serial Correlation in the Residuals
6. Forecast VAR

1. Visualizing the Data: uschange data (Revisted)

```
us_change |>
  pivot_longer(!Quarter,
               names_to = "Series") |>
  ggplot(aes(x = Quarter, y = value,
             col = Series)) +
  geom_line() +
  facet_wrap(.~Series, scales = "free_y") +
  theme(legend.position = "none")
```



Consumption
Income
Production
Savings
Unemployment
value
Quarter
1980 Q1
2000 Q1
2020 Q1

`uschange` data (Revisted)

- It is logical that the relationship between these variables is not contemporaneous as we assumed earlier in the semester.
- Instead, last quarter's personal disposable income could affect this quarter's consumption and so forth.
- We can therefore use a VAR to determine the relationship between these series.

2. Check for Stationarity?

Visually:

- ACF
- autoplot

Formally:

- kpss test

Conclusion

These growth data are potentially stationary. We will assume they are for the purposes of this exercise.

2. Check for Stationarity

- If the series are stationary, we forecast them by fitting a VAR to the data directly (known as a "VAR in levels").
- If the series are non-stationary, we can take differences of the data in order to make them stationary, then fit a VAR model (known as a "VAR in differences").
- In both cases, the models are estimated equation by equation using the principle of least squares.
    - For each equation, the parameters are estimated by minimizing the sum of squared $\varepsilon_{i,t}$ values.

3. Select the order (p) of the VAR Model

The optimal lag length can be selected by iteratively fitting increasing orders of VAR model and pick the order that gives a model with least AIC or BIC value.

```
#require(vars)
vars::VARselect(us_change[,2:3], lag.max = 10, type = c("const
season = NULL, exogen = NULL)
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      5      1      1      5
##
## $criteria
##                  1          2          3          4
## AIC(n) -1.4959327 -1.4863240 -1.5277019 -1.5398494 -1.57535
## HQ(n)  -1.4540832 -1.4165748 -1.4300531 -1.4143009 -1.42191
## SC(n)  -1.3926420 -1.3141728 -1.2866903 -1.2299773 -1.19662
## FPE(n)  0.2240408  0.2262083  0.2170488  0.2144448  0.20698
##                  7         8          9         10
## AIC(n) -1.5282562 -1.510674 -1.4720448 -1.4407631
## HQ(n)  -1.3190087 -1.273527 -1.2069979 -1.1478166
## SC(n)  -1.0118027 -0.925360 -0.8178704 -0.7177282
## FPE(n)  0.2170611  0.220980  0.2297738  0.2371907
```

Lag Selection

The AIC criteria selects a model with 5 lags and the BIC chooses a more parsimonious model, with 1 lag.

```r
us_change |>
  model(
    aicc = VAR(vars(Consumption, Income)),
    aic = VAR(vars(Consumption, Income),
               ic = "aic"),
    bic = VAR(vars(Consumption, Income),
              ic = "bic")
  ) -> fit.VAR

fit.VAR
```

```
## # A mable: 1 x 3
##                aicc                 aic                 bic
##             <model>             <model>             <model>
## 1 <VAR(5) w/ mean> <VAR(5) w/ mean> <VAR(1) w/ mean>
```

Caution!!!

Care should be taken when using the AIC/AICc as they tend to choose large numbers of lags.

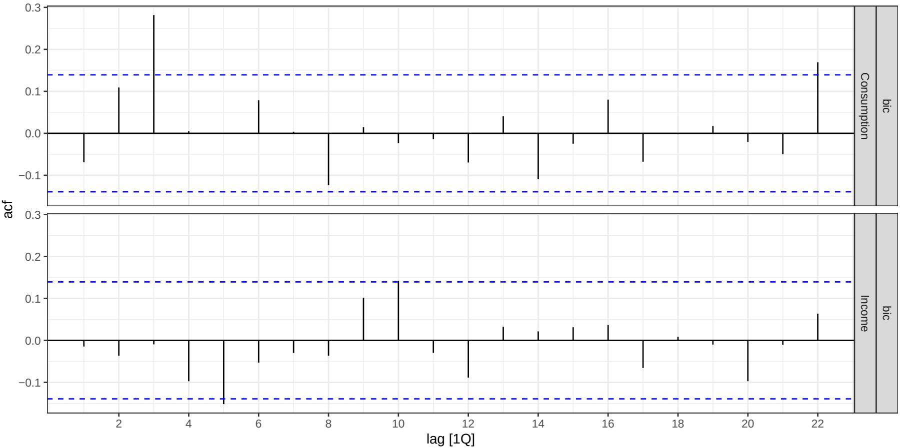
```
fit.VAR |> glance()
```

```
## # A tibble: 3 x 6
##   .model sigma2          log_lik   AIC  AICc   BIC
##   <chr>  <list>            <dbl> <dbl> <dbl> <dbl>
## 1 aicc   <dbl [2 x 2]>    -373.  798.  806.  883.
## 2 aic    <dbl [2 x 2]>    -373.  798.  806.  883.
## 3 bic    <dbl [2 x 2]>    -408.  836.  837.  869.
```

4 & 5. Evaluating the implied models

Our contending model results are stored in `fit.VAR` above. Let us start with the more parsimonious model, VAR(1), as implied by the BIC criteria.

```
fit.VAR |> select(bic) |> augment() |>
  ACF(.innov) |> autoplot()
```



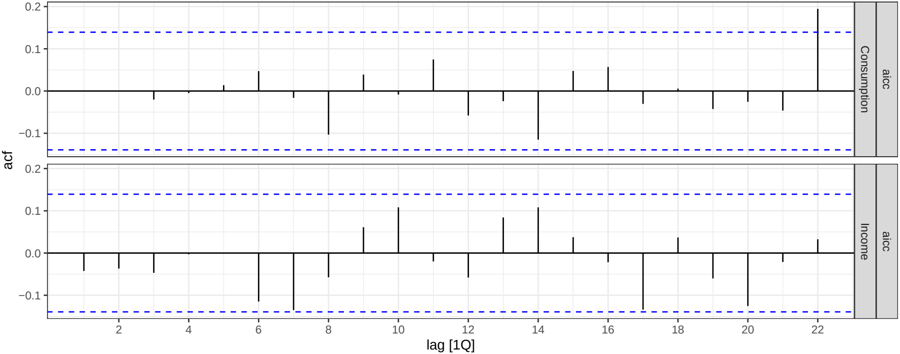
Observations

There appears to be autocorrelation remaining in the Consumption series from the VAR(1) model.

4 & 5. Evaluating the implied models

Turning to the VAR(5) as selected by the AICc.

```
fit.VAR |> select(aicc) |> augment() |>
  ACF(.innov) |> autoplot()
```



Observations

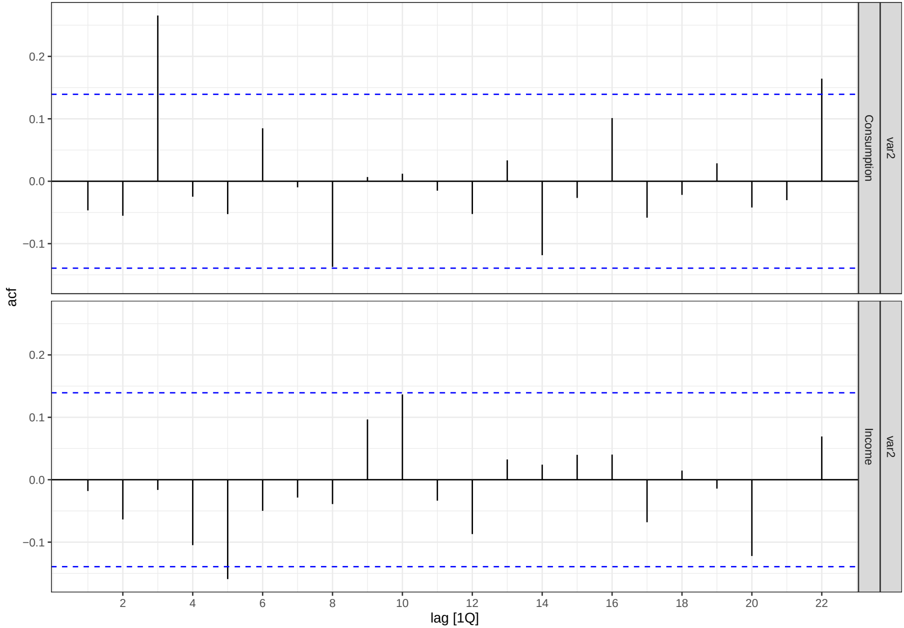
The VAR(5) model appears to adequately capture the Consumption data as there is no feature remaining in the model innovations.

4 & 5. Evaluating the implied models

In practice, we might settle somewhere in between the two models. For example, we could reestimate a VAR(2) then check if that passes the diagnostic tests. If it doesn't, we move on to a VAR(3), rinse and repeat.

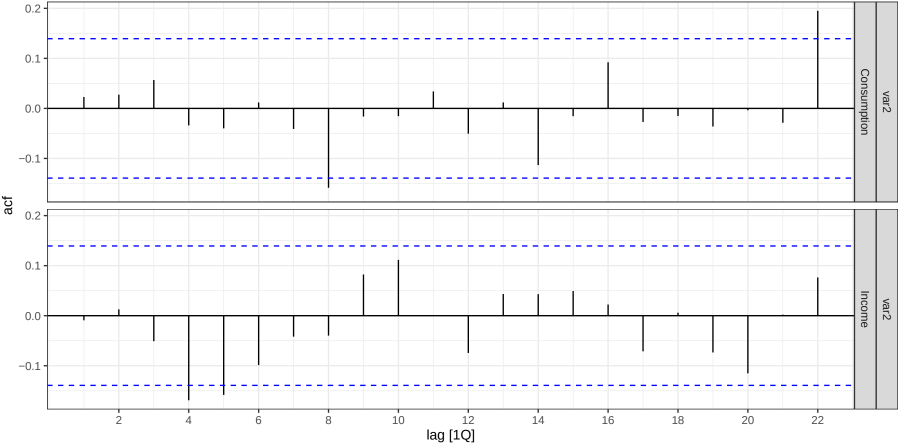
```
var2 <- us_change |>
   model(var2 = VAR(
    vars(Consumption, Income) ~ AR(2))
  )
```

```
var2 |> augment() |> ACF(.innov) |> autoplot()
```



```r
var3 <- us_change |>
   model(var2 = VAR(
    vars(Consumption, Income) ~ AR(3))
  )

var3 |> augment() |> ACF(.innov) |> autoplot()
```



Final Remarks

It appears that we could proceed with the VAR(3) model after all.

Estimation by Least Squares

Creating the first order lags of the variables:

```
us_change2 <- us_change |>
  select(Quarter, Consumption, Income) |>
  mutate(
    l.Consumption = Consumption |> lag(1),
    l.Income = Income |> lag(1),
         )

us_change2 |> head()
```

```
## # A tsibble: 6 x 5 [1Q]
##   Quarter Consumption Income l.Consumption l.Income
##     <qtr>       <dbl>  <dbl>         <dbl>    <dbl>
## 1 1970 Q1       0.619  1.04             NA       NA
## 2 1970 Q2       0.452  1.23          0.619     1.04
## 3 1970 Q3       0.873  1.59          0.452     1.23
## 4 1970 Q4      -0.272 -0.240         0.873     1.59
```

Estimation by Least Squares vs VAR

Estimate the LM equations:

```
# Consumption Equation
con.ols <- us_change2 |> model(
  con = TSLM(Consumption ~ l.Consumption + l.Income))

# Income Equation
inc.ols <- us_change2 |> model(
  con = TSLM(Income ~ l.Consumption + l.Income))
```

```
con.ols |> report()
```

```
## Series: Consumption
## Model: TSLM
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -2.743211 -0.354520  0.003287  0.342220  1.536232
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.46429    0.06767   6.862 8.92e-11 ***
## l.Consumption  0.26530    0.07267   3.651 0.000336 ***
## l.Income       0.11297    0.05144   2.196 0.029277 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
##
## Residual standard error: 0.6001 on 194 degrees of freedom
## Multiple R-squared: 0.1282,  Adjusted R-squared: 0.1192
## F-statistic: 14.27 on 2 and 194 DF, p-value: 1.6584e-06
```

```r
#VAR(1) model coefficients
fit.VAR |> select(bic) |> report()
```

```
## Series: Consumption, Income
## Model: VAR(1) w/ mean
##
## Coefficients for Consumption:
##       lag(Consumption,1)  lag(Income,1)  constant
##                   0.2653         0.1130    0.4643
## s.e.              0.0727         0.0514    0.0677
##
## Coefficients for Income:
##       lag(Consumption,1)  lag(Income,1)  constant
##                   0.4919        -0.2576     0.550
## s.e.              0.1031         0.0730     0.096
##
## Residual covariance matrix:
##             Consumption Income
## Consumption      0.3601 0.1970
## Income           0.1970 0.7251
##
## log likelihood = -407.93
## AIC = 835.86 AICc = 837.04   BIC = 868.69
```

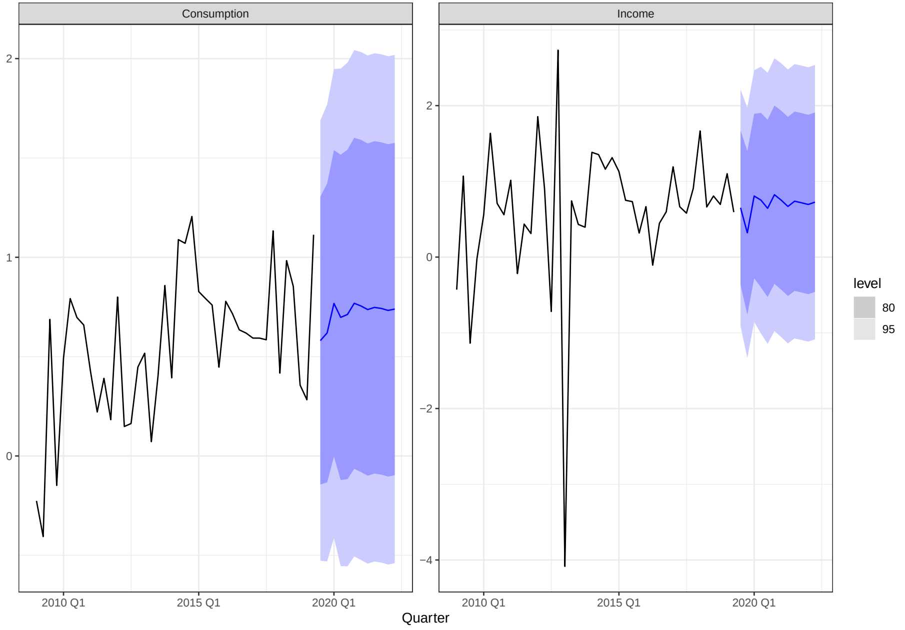
```r
#VAR(1) model coefficients
fit.VAR |> select(bic) |> tidy() |>
  filter(.response == "Consumption")|>
  select(-.model, -.response)
```

```
## # A tibble: 3 x 5
##   term                 estimate std.error statistic  p.value
##   <chr>                   <dbl>     <dbl>     <dbl>    <dbl>
## 1 lag(Consumption,1)      0.265    0.0727      3.65 3.36e- 4
## 2 lag(Income,1)           0.113    0.0514      2.20 2.93e- 2
## 3 constant                0.464    0.0677      6.86 8.92e-11
```

6. Forecast VAR

$$\hat{y}_{1,T+1|T} = \hat{\gamma}_1 + \hat{\phi}_{11} y_{1,T} + \hat{\phi}_{12} y_{2,T}$$
$$\hat{y}_{2,T+1|T} = \hat{\gamma}_2 + \hat{\phi}_{21} y_{1,T} + \hat{\phi}_{22} y_{2,T}.$$

```
fit.VAR |> select(aicc)  |>  forecast(h = "3years")  |>
    autoplot(us_change |>
  filter(year(Quarter) > 2008))
```


Consumption
Income
2
1
0
−2
−4
2010 Q1
2015 Q1
2020 Q1
Quarter
level
80
95

Remarks

A criticism that VARs face is that they are atheoretical; that is, they are not built on some economic theory that imposes a theoretical structure on the equations. Every variable is assumed to influence every other variable in the system, which makes a direct interpretation of the estimated coefficients difficult.

Despite this, VARs are useful in several contexts (see earlier slides).

Section 2

Impulse Response Functions (IRF)

Impulse Response Functions (IRF)

Coming Soon!!!

Coming to a GitHub Page Near you!

Section 3

Forecast Error Variance Decomposition (FEVD)

Forecast Error Variance Decomposition (FEVD)

Coming Soon!!!

Coming to a GitHub Page Near you!