

# Séries Temporelles Multivariées

## Chapitre 4

### Économétrie des séries temporelles





# Structure du cours :

Autorégressions Vectorielles – VAR

Exemples de base

Propriétés

- ▶ Stationnarité

Revisiter les processus ARMA univariés

Prévision

- ▶ Causalité de Granger
- ▶ Fonctions de réponse impulsionnelle

Cointégration

- ▶ Examiner les relations à long terme
- ▶ Déterminer si un VAR est cointégré
- ▶ Modèles de Correction d'Erreur
- ▶ Tests de Cointégration - Engle-Granger

## Autorégressions Vectorielles – VAR

# Pourquoi l'analyse VAR ?

## VAR Stationnaires

- ▶ Déterminer si les variables se rétroagissent les unes sur les autres
- ▶ Améliorer les prévisions
- ▶ Modéliser l'effet d'un choc dans une série sur une autre
- ▶ Différencier entre les dynamiques à court terme et à long terme

## Cointégration

- ▶ Lier les marches aléatoires
- ▶ Découvrir les relations à long terme
- ▶ Peut améliorer considérablement les prévisions à moyen et long terme

## Définition du VAR

**Autorégression d'ordre P, AR(P) :**

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_P y_{t-P} + \varepsilon_t$$

**Autorégression vectorielle d'ordre P, VAR(P) :**

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t$$

où  $\mathbf{y}_t$  et  $\varepsilon_t$  sont des vecteurs de dimension  $k \times 1$

**VAR(1) bivarié :**

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{01} \\ \phi_{02} \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

**Exprime de manière compacte deux modèles liés :**

$$y_{1,t} = \phi_{01} + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = \phi_{02} + \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + \varepsilon_{2,t}$$

La stationnarité est une forme statistiquement significative de régularité.

Un processus stochastique  $y_t$  est stationnaire en covariance si :

- ▶  $E[y_t] = \mu, \quad \forall t$
- ▶  $V[y_t] = \sigma^2, \quad \sigma^2 < \infty, \quad \forall t$
- ▶  $E[(y_t - \mu)(y_{t-s} - \mu)] = \gamma_s, \quad \forall t, s$

**Stationnarité de l'AR(1) :**  $y_t = \phi y_{t-1} + \varepsilon_t$

- ▶  $|\phi| < 1$
- ▶  $\varepsilon_t$  est un bruit blanc

**Stationnarité de l'AR(P) :**  $y_t = \phi_1 y_{t-1} + \dots + \phi_P y_{t-P} + \varepsilon_t$

- ▶ Les racines de  $(z^P - \phi_1 z^{P-1} - \phi_2 z^{P-2} - \dots - \phi_{P-1} z - \phi_P)$  sont inférieures à 1
- ▶  $\varepsilon_t$  est un bruit blanc

*Pas de dépendance à  $t$*



### AR(1)

$$\begin{aligned}y_t &= \phi_0 + \phi_1 y_{t-1} + \varepsilon_t \\&= \phi_0 + \phi_1 (\phi_0 + \phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2 (\phi_0 + \phi_1 y_{t-3} + \varepsilon_{t-2}) + \phi_1 \varepsilon_{t-1} + \varepsilon_t \\&\vdots \\&= \sum_{i=0}^{\infty} \phi_1^i \phi_0 + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i} \\&= (1 - \phi_1)^{-1} \phi_0 + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}\end{aligned}$$

### VAR(1)

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 (\boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_{t-1}) + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1^2 \mathbf{y}_{t-2} + \boldsymbol{\Phi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1^2 (\boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-3} + \boldsymbol{\varepsilon}_{t-2}) + \boldsymbol{\Phi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t \\ &\vdots \\ &= \sum_{i=0}^{\infty} \boldsymbol{\Phi}_1^i \boldsymbol{\Phi}_0 + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_1^i \boldsymbol{\varepsilon}_{t-i} \\ &= (\mathbf{I}_k - \boldsymbol{\Phi}_1)^{-1} \boldsymbol{\Phi}_0 + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_1^i \boldsymbol{\varepsilon}_{t-i} \end{aligned}$$

# Propriétés d'un VAR(1) et AR(1)

## Moyenne, Variance, Autocovariance

Propriété	AR(1)	VAR(1)
Équation	$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$	$\mathbf{y}_t = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \varepsilon_t$
Moyenne	$\frac{\phi_0}{1-\phi_1}$	$(\mathbf{I}_k - \boldsymbol{\Phi}_1)^{-1} \boldsymbol{\Phi}_0$
Variance	$\frac{\sigma^2}{1-\phi_1^2}$	$(\mathbf{I} - \boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)^{-1} \text{vec}(\boldsymbol{\Sigma})$
Autocovariance $s$	$\gamma_s = \phi_1^s V[y_t]$	$\boldsymbol{\Gamma}_s = \boldsymbol{\Phi}_1^s V[\mathbf{y}_t]$
Autocovariance $-s$	$\gamma_{-s} = \phi_1^s V[y_t]$	$\boldsymbol{\Gamma}_{-s} = V[\mathbf{y}_t] \boldsymbol{\Phi}_1^{s'}$

Les autocovariances des processus vectoriels ne sont pas symétriques, mais  $\boldsymbol{\Gamma}_s = \boldsymbol{\Gamma}_{-s}'$

## Stationnarité

- ▶ AR(1) :  $|\phi_1| < 1$
- ▶ VAR(1) :  $|\lambda_i| < 1$  où  $\lambda_i$  sont les valeurs propres de  $\boldsymbol{\Phi}_1$

### Modèle

$$\begin{pmatrix} \text{PIB}_t \\ \text{CHO}_t \end{pmatrix} = \begin{pmatrix} 0.547^{***} \\ 0.049^* \end{pmatrix} + \begin{pmatrix} 0.182^{**} & -0.639^{***} \\ -0.096^{***} & 0.507^{***} \end{pmatrix} \begin{pmatrix} \text{PIB}_{t-1} \\ \text{CHO}_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

Significance : \*\*\* : 1%, \*\* : 5%, \* : 10%

## Modèle du PIB

$$\text{PIB}_t = \phi_{01} + \phi_{11,1}\text{PIB}_{t-1} + \phi_{12,1}\text{CHO}_{t-1} + \varepsilon_{1,t}$$

## Modèle du Chomage

$$\text{CHO}_t = \phi_{02} + \phi_{21,1}\text{PIB}_{t-1} + \phi_{22,1}\text{CHO}_{t-1} + \varepsilon_{2,t}$$

## Estimations VAR(1)

$$\text{PIB}_t = 0.547^{***} + 0.182^{**}\text{PIB}_{t-1} - 0.639^{***}\text{CHO}_{t-1} + \varepsilon_{1,t}$$

$$\text{CHO}_t = 0.049^* - 0.096^{***}\text{PIB}_{t-1} + 0.507^{***}\text{CHO}_{t-1} + \varepsilon_{2,t}$$

## Estimations AR(1)

$$\text{PIB}_t = 0.648^{***} + 0.414^{***}\text{PIB}_{t-1} + \varepsilon_{1,t}$$

$$\text{CHO}_t = -0.003 + 0.692^{***}\text{CHO}_{t-1} + \varepsilon_{2,t}$$

# Données

$y_1$  : **PIB**

$y_2$  : **Chomâge**

Pays : **USA**

Période : **jan. 1980 - déc. 2019**

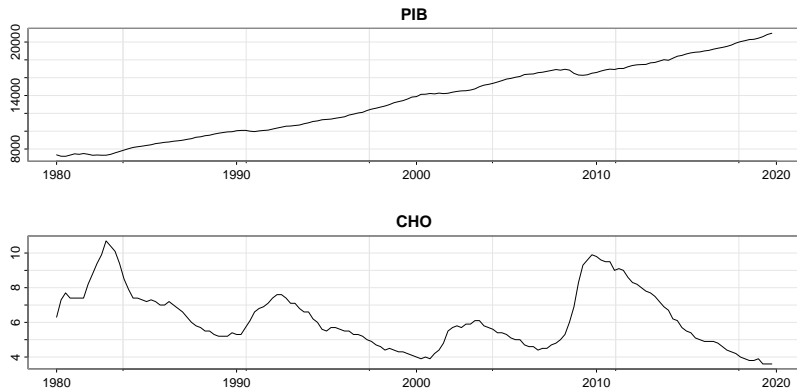
Fréquence : **trimestrielle**

## Les données – Obtention

```
start_date <- as.Date("1980-01-01")
#start_date <- as.Date("1957-01-01")
end_date <- as.Date("2019-12-31")
#end_date <- as.Date("2024-12-31")

# 1. Télécharger et transformer les données FRED
gdp_data <- fredr(series_id = "GDPC1",
                  observation_start = start_date,
                  observation_end = end_date, frequency = "q")
unemp_data <- fredr(series_id = "UNRATE",
                    observation_start = start_date,
                    observation_end = end_date, frequency = "q")
```

## Graphs – données brutes





## Les données – Manipulation

```
# merge les deux variables dans une data.frame
data <- merge(gdp_data[, c("date", "value")],
              unemp_data[, c("date", "value")],
              by = "date", suffixes = c("_gdp", "_unemp"))
colnames(data) <- c("date", "gdp", "unemp")

# transformation des variables
data <- data %>%
  mutate(log_gdp = log(gdp), unemp = unemp) %>%
  select(date, log_gdp, unemp)

# s'assurer que les données sont stationnaires (différencier si nécessaire)
data_diff <- data %>%
  mutate(d_log_gdp = c(NA, diff(log_gdp)),
         d_unemp = c(NA, diff(unemp))) %>%
  na.omit()

# autre méthode pour transformer les données (à faire).
```

## Les données – Tester la stationarité

**PIB – log et diff(log)**

```
tseries::adf.test(data$log_gdp)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: data$log_gdp  
## Dickey-Fuller = -1.3962, Lag order = 5, p-value = 0.8282  
## alternative hypothesis: stationary
```

```
tseries::adf.test(data_diff$d_log_gdp)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: data_diff$d_log_gdp  
## Dickey-Fuller = -4.3656, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

## Les données – Tester la stationarité

**CHO** – taux et variation de taux

```
tseries::adf.test(data$unemp)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: data$unemp  
## Dickey-Fuller = -2.5066, Lag order = 5, p-value = 0.3651  
## alternative hypothesis: stationary
```

```
tseries::adf.test(data_diff$d_unemp)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: data_diff$d_unemp  
## Dickey-Fuller = -3.8569, Lag order = 5, p-value = 0.01802  
## alternative hypothesis: stationary
```

## Les modèles – Estimation VAR(1) vs AR(1)

```
# 2. Estimation des modèles
# (a) Modèle VAR(1) multivarié
data_diff$d_log_gdp <- data_diff$d_log_gdp*100
# *100 pour être en % et coller à l'échelle de la variation du taux de chômage
var_model <- vars::VAR(data_diff[, c("d_log_gdp", "d_unemp")],
                       p = 1, type = "const")

# (b) Modèles AR(1) univariés
ar_gdp <- arima(data_diff$d_log_gdp, order = c(1, 0, 0))
ar_unemp <- arima(data_diff$d_unemp, order = c(1, 0, 0))
```

## VAR(1) – Résultats d'estimation

```
# 3. Analyse des résultats
# Résultats du VAR(1)

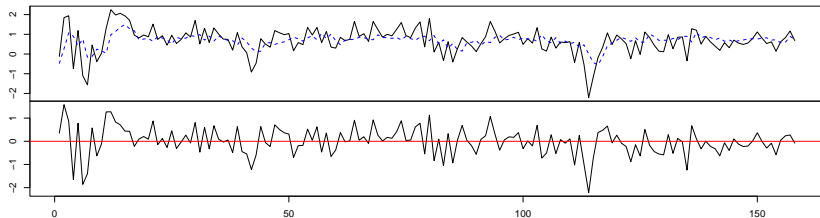
# lancer tout de meme summary dans le rmd
# summary(var_model)

coeftest(var_model)
```

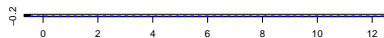
```
##
## t test of coefficients:
##
##              Estimate Std. Error t value  Pr(>|t|)
## d_log_gdp:(Intercept)  0.546785   0.074933  7.2970 1.430e-11 ***
## d_log_gdp:d_log_gdp.l1  0.182007   0.091499  1.9892 0.048442 *
## d_log_gdp:d_unemp.l1   -0.639115   0.202632 -3.1541 0.001934 **
## d_unemp:(Intercept)    0.048835   0.027895  1.7507 0.081981 .
## d_unemp:d_log_gdp.l1   -0.096275   0.034062 -2.8265 0.005327 **
## d_unemp:d_unemp.l1     0.507314   0.075432  6.7254 3.178e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Graph des résidus

Diagram of fit and residuals for d\_log\_gdp



ACF Residuals



PACF Residuals

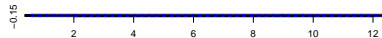
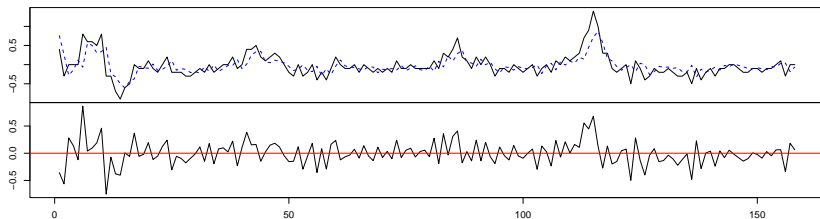


Diagram of fit and residuals for d\_unemp



ACF Residuals



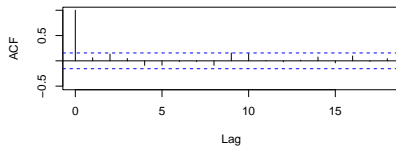
PACF Residuals



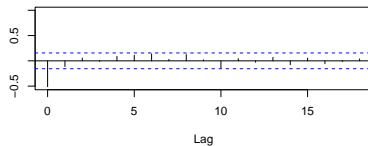
# ACF Résidus

```
acf(residuals(var_model))
```

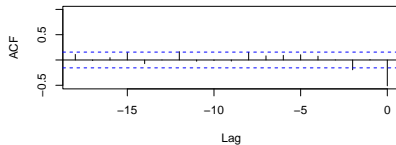
**d\_log\_gdp**



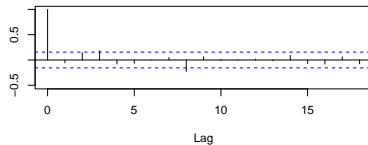
**d\_log\_gdp & d\_unemp**



**d\_unemp & d\_log\_gdp**

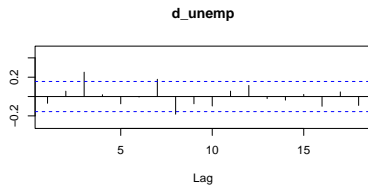
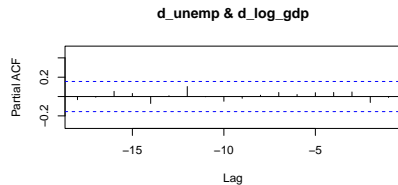
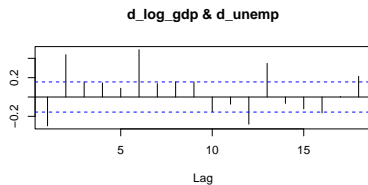
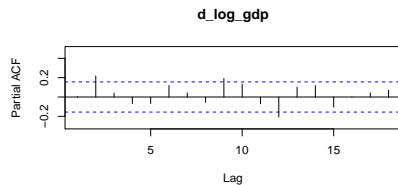


**d\_unemp**



# PACF Résidus

```
pacf(residuals(var_model))
```





# AR(1) – Résultats d'estimation

```
# Résultats des modèles AR(1)
```

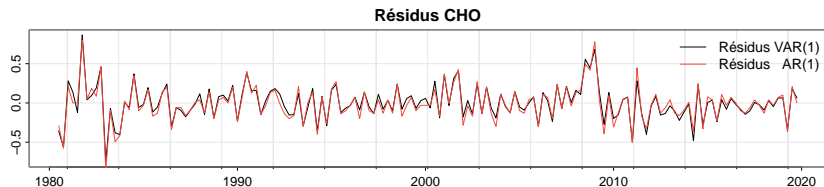
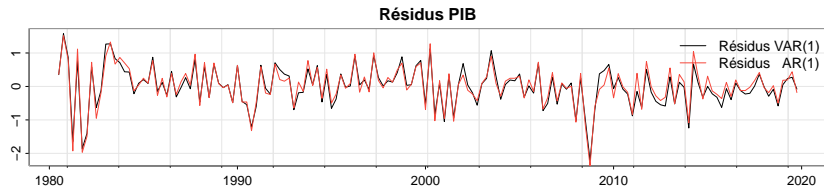
```
coeftest(ar_gdp)
```

```
##
## z test of coefficients:
##
##           Estimate Std. Error z value Pr(>|z|)
## ar1          0.413760   0.076315  5.4217 5.903e-08 ***
## intercept    0.648535   0.085898  7.5501 4.350e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coeftest(ar_unemp)
```

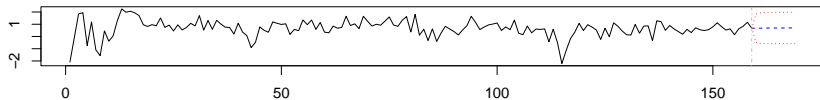
```
##
## z test of coefficients:
##
##           Estimate Std. Error z value Pr(>|z|)
## ar1          0.6918376  0.0600663 11.5179 <2e-16 ***
## intercept   -0.0027798  0.0589276 -0.0472  0.9624
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Résidus VAR(1) vs AR(1)

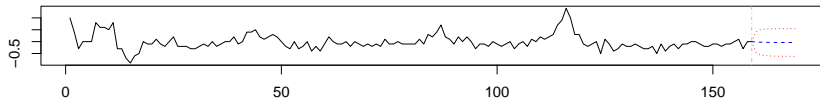


# Prévision VAR(1)

**Forecast of series d\_log\_gdp**

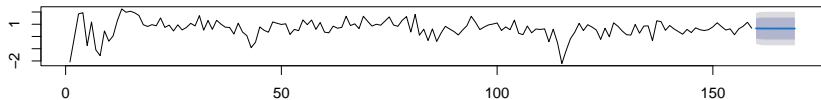


**Forecast of series d\_unemp**

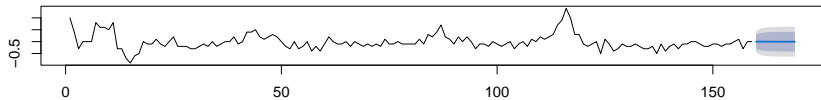


# Prévision AR(1)

**Forecasts from ARIMA(1,0,0) with non-zero mean**



**Forecasts from ARIMA(1,0,0) with non-zero mean**



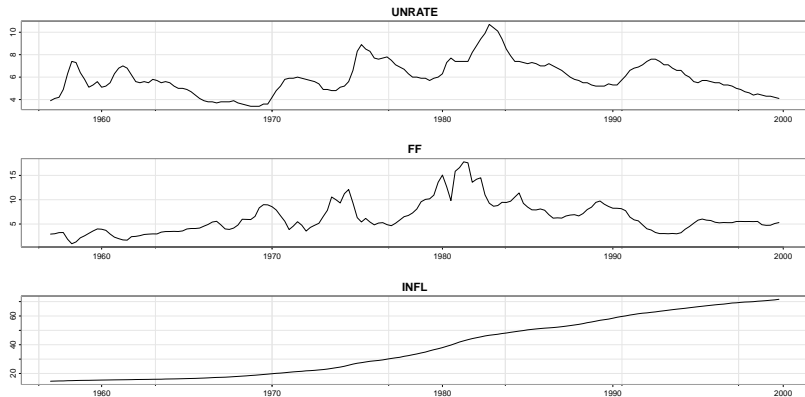
## ~ Taylor Rule : Chomage, Fed Funds, Inflation

$y_1$  : **Chomage**,  $y_2$  : **Fed Funds**  $y_3$  : **Inflation** Pays : **USA** Période : **jan. 1957 - déc. 2019** Fréquence : **trimestrielle**

```
start_date <- as.Date("1957-01-01")
end_date <- as.Date("1999-12-31")
# start_date <- as.Date("1980-01-01")
# end_date <- as.Date("2019-12-31")

UNRATE <- fredr(series_id = "UNRATE",
                observation_start = start_date,
                observation_end = end_date,
                frequency = "q")
DFF <- fredr(series_id = "DFF",
             observation_start = start_date,
             observation_end = end_date,
             frequency = "q")
GDPDEF <- fredr(series_id = "GDPDEF",
                observation_start = start_date,
                observation_end = end_date,
                frequency = "q")
```

## Stationarité – visuelle



## Stationarité – test

```
adf_unrate <- tseries::adf.test(UNRATE$value)
adf_dff <- tseries::adf.test(DFF$value)
adf_gdpdef <- tseries::adf.test(GDPDEF$value)

adf_values <- data.frame(
  Série = c("UNRATE", "DFF", "GDPDEF"),
  `Statistique ADF` = c(adf_unrate$statistic,
                        adf_dff$statistic,
                        adf_gdpdef$statistic),
  `p-value` = c(adf_unrate$p.value,
                adf_dff$p.value,
                adf_gdpdef$p.value)
)
adf_values
```

```
##      Série Statistique.ADF    p.value
## 1 UNRATE      -2.355513 0.42780874
## 2   DFF       -3.150157 0.09839603
## 3 GDPDEF      -2.630979 0.31277964
```

## Différenciation des séries

```
# on différencie avec des méthodes qui maintiennent
# une "cohérence" des nouvelles variables
UNRATE.d <- na.omit(diff(UNRATE$value))
DFF.d <- na.omit(diff(DFF$value))
GDPDEF.d <- na.omit((log(diff(GDPDEF$value))))
# test stationarité
adf_unrate_d <- tseries::adf.test(UNRATE.d)
adf_dff_d <- tseries::adf.test(DFF.d)
adf_gdpdef_d <- tseries::adf.test(GDPDEF.d)

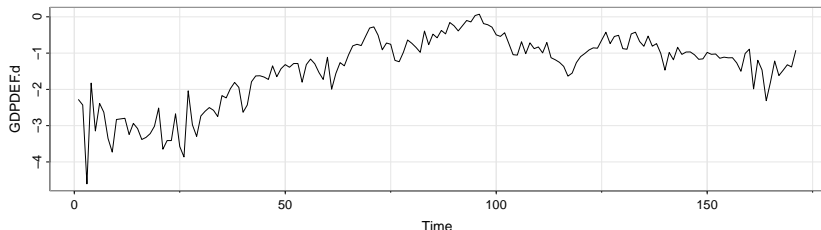
adf_values_d <- data.frame(
  Série = c("UNRATE", "DFF", "GDPDEF"),
  `Statistique ADF` = c(adf_unrate_d$statistic,
                        adf_dff_d$statistic,
                        adf_gdpdef_d$statistic),
  `p-value` = c(adf_unrate_d$p.value,
                adf_dff_d$p.value,
                adf_gdpdef_d$p.value)
)
adf_values_d
```

```
##      Série Statistique.ADF    p.value
## 1 UNRATE      -5.218266 0.0100000
## 2   DFF      -5.207432 0.0100000
## 3 GDPDEF      -1.165658 0.9093566
```



## Graph de la série $I(2) \rightarrow I(1) \rightarrow I(0)$

```
tsplot(GDPDEF.d)
```

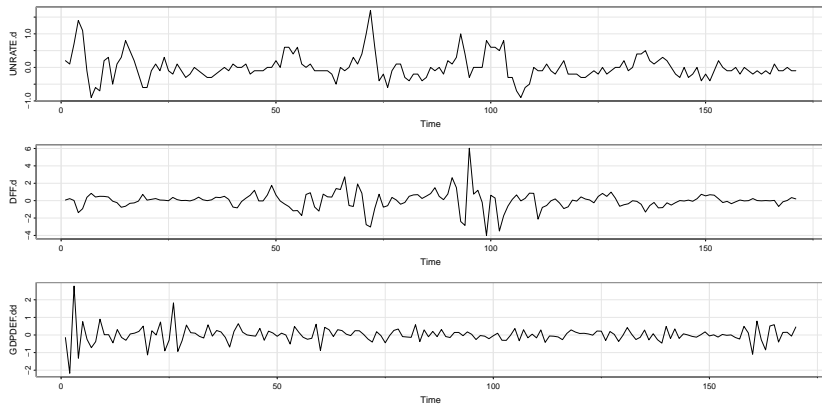


```
# GDPDEF.d n'est toujours pas stationnaire : on différencie une seconde fois  
GDPDEF.dd <- na.omit(diff((log(diff(GDPDEF$value))))))  
# test stationarité  
tseries::adf.test(diff(GDPDEF.dd))
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: diff(GDPDEF.dd)  
## Dickey-Fuller = -10.4, Lag order = 5, p-value = 0.01  
## alternative hypothesis: stationary
```

# Graph des séries stationnaires

```
par(mfrow=c(3,1))  
tsplot(UNRATE.d)  
tsplot(DFF.d)  
tsplot(GDPDEF.dd)
```



## Estimation VAR(1) – UNRATE, FF, INF

```
PolMon <- VAR(cbind(UNRATE.d, DFF.d, GDPDEF.dd), type = "none")
```

```
#summary(PolMon)
```

```
coef(PolMon)
```

```
## $UNRATE.d
```

	Estimate	Std. Error	t value	Pr(> t )
## UNRATE.d.l1	0.65596501	0.06919730	9.479633	2.512014e-17
## DFF.d.l1	0.03234603	0.02530148	1.278424	2.028737e-01
## GDPDEF.dd.l1	0.05270480	0.04886464	1.078588	2.823270e-01

```
##
```

```
## $DFF.d
```

	Estimate	Std. Error	t value	Pr(> t )
## UNRATE.d.l1	-0.79875010	0.22816004	-3.5008325	0.0005949147
## DFF.d.l1	0.06938984	0.08342504	0.8317628	0.4067305799
## GDPDEF.dd.l1	0.02928980	0.16111840	0.1817905	0.8559675801

```
##
```

```
## $GDPDEF.dd
```

	Estimate	Std. Error	t value	Pr(> t )
## UNRATE.d.l1	-0.055376927	0.09655546	-0.5735245	5.670611e-01
## DFF.d.l1	0.003746848	0.03530479	0.1061286	9.156077e-01
## GDPDEF.dd.l1	-0.479475102	0.06818399	-7.0320779	4.973776e-11

## Modèle trivarié VAR(1) estimé

$$UNRATE_t = 0.656^{***} UNRATE_{t-1} + 0.032FF_{t-1} + 0.053INF_{t-1} + \varepsilon_{UNRATE,t}$$

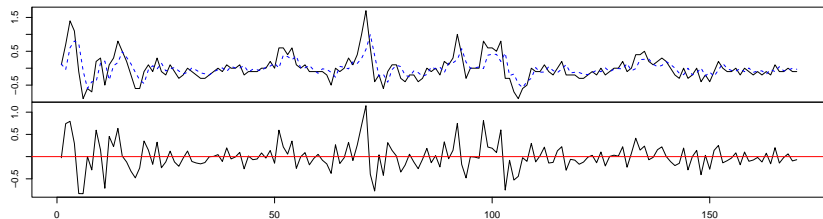
$$FF_t = -0.799^{***} UNRATE_{t-1} + 0.069FF_{t-1} + 0.029INF_{t-1} + \varepsilon_{FF,t}$$

$$INF_t = -0.055UNRATE_{t-1} + 0.004FF_{t-1} - 0.479^{***} INF_{t-1} + \varepsilon_{INF,t}$$

*En pratique, pour vérifier la Taylor Rule on applique un VAR structurel (SVAR) afin d'imposer une relation **structurelle***

# Graph $y_t, \hat{y}_t, \hat{\varepsilon}_t, ACF$ et $PACF$ de $\hat{\varepsilon}_t$

Diagram of fit and residuals for UNRATE.d

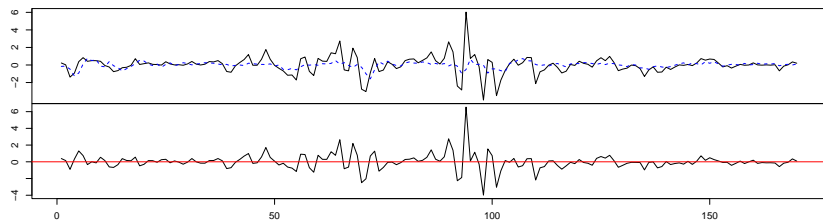


ACF Residuals

PACF Residuals



Diagram of fit and residuals for DFF.d



ACF Residuals

PACF Residuals



VAR(P) est en réalité un VAR(1)

Forme compagnon :

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t$$

Reformuler en un seul VAR(1) où  $\mu = E[\mathbf{y}_t] = (I - \Phi_1 - \dots - \Phi_P)^{-1} \Phi_0$

$$\mathbf{z}_t = \Upsilon \mathbf{z}_{t-1} + \xi_t$$

$$\mathbf{z}_t = \begin{pmatrix} \mathbf{y}_t - \mu \\ \mathbf{y}_{t-1} - \mu \\ \vdots \\ \mathbf{y}_{t-P+1} - \mu \end{pmatrix}, \Upsilon = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \dots & \Phi_{P-1} & \Phi_P \\ I_k & 0 & 0 & \dots & 0 & 0 \\ 0 & I_k & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I_k & 0 \end{pmatrix}$$

- Tous les résultats peuvent être directement appliqués à la forme compagnon.
- Peut également être utilisé pour transformer AR(P) en VAR(1)

## Revisiter les prévisions univariées

- Considérons l'AR(1) standard

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

- Prédiction optimale à 1 étape :

$$\begin{aligned} E_t [y_{t+1}] &= E_t [\phi_0] + E_t [\phi_1 y_t] + E_t [\varepsilon_{t+1}] \\ &= \phi_0 + \phi_1 y_t + 0 \end{aligned}$$

- Prédiction optimale à 2 étapes :

$$\begin{aligned} E_t [y_{t+2}] &= E_t [\phi_0] + E_t [\phi_1 y_{t+1}] + E_t [\varepsilon_{t+2}] \\ &= \phi_0 + \phi_1 E_t [y_{t+1}] + 0 \\ &= \phi_0 + \phi_1 (\phi_0 + \phi_1 y_t) \\ &= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_t \end{aligned}$$

- Prédiction optimale à  $h$  étapes :

$$E_t [y_{t+h}] = \sum_{i=0}^{h-1} \phi_1^i \phi_0 + \phi_1^h y_t$$

## Prévisions avec VAR

- Identique au cas univarié

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \varepsilon_t$$

- Prédiction optimale à 1 étape :

$$\begin{aligned} E_t [\mathbf{y}_{t+1}] &= E_t [\Phi_0] + E_t [\Phi_1 \mathbf{y}_t] + E_t [\varepsilon_{t+1}] \\ &= \Phi_0 + \Phi_1 \mathbf{y}_t + \mathbf{0} \end{aligned}$$

- Prédiction optimale à  $h$  étapes :

$$\begin{aligned} E_t [\mathbf{y}_{t+h}] &= \Phi_0 + \Phi_1 \Phi_0 + \dots + \Phi_1^{h-1} \Phi_0 + \Phi_1^h \mathbf{y}_t \\ &= \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t \end{aligned}$$

- Prédiction de plus haut ordre peut être calculée de manière récursive

$$E_t [\mathbf{y}_{t+h}] = \Phi_0 + \Phi_1 E_t [\mathbf{y}_{t+h-1}] + \dots + \Phi_P E_t [\mathbf{y}_{t+h-P}]$$



## Qu'est-ce qui fait une bonne prévision ?

- ▶ Résidus de prévision

$$\hat{e}_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$$

- ▶ Les résidus *ne sont pas* un bruit blanc
- ▶ Peut contenir une composante  $MA(h-1)$ 
  - ▶ Erreur de prévision pour  $y_{t+1} - \hat{y}_{t+1|t-h+1}$  n'était pas connue au moment  $t$ .
- ▶ Tracez vos résidus
- ▶ ACF des résidus
- ▶ Régressions de Mincer-Zarnowitz
- ▶ Procédure à trois périodes
  - ▶ Échantillon d'entraînement : Utilisé pour construire le modèle
  - ▶ Échantillon de validation : Utilisé pour affiner le modèle
  - ▶ Échantillon d'évaluation : Test ultime, idéalement en une seule fois

## Prévision multi-étapes

### Deux méthodes

#### ► Méthode itérative

- Construire un modèle pour les prévisions à 1 étape

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \varepsilon_t$$

- Itérer la prévision jusqu'à la période  $h$

$$\hat{\mathbf{y}}_{t+h|t} = \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t$$

- Utilise efficacement les informations
- Impose beaucoup de structure au problème

#### ► Méthode directe

- Construire un modèle pour les prévisions à  $h$  étapes

$$\mathbf{y}_t = \Phi_0 + \Phi_h \mathbf{y}_{t-h} + \varepsilon_t$$

- Prévision directe en utilisant une méthode pseudo à 1 étape

$$\hat{\mathbf{y}}_{t+h|t} = \Phi_0 + \Phi_h \mathbf{y}_t$$

- Robuste à certaines non-linéarités

## Évaluation des prévisions multi-étapes

- ▶ L'évaluation des prévisions multi-étapes est identique à l'évaluation des prévisions à une étape avec une réserve
- ▶ Les erreurs de prévision à  $h$  étapes peuvent être corrélées avec toute erreur de prévision non connue au moment  $t$

$$\hat{e}_{t+1|t-h+1}, \hat{e}_{t+2|t-h+2}, \dots, \hat{e}_{t+h-1|t-1}$$

- ▶ Conduit à une structure  $MA(h-1)$  dans les erreurs de prévision
- ▶ Solutions :
  - ▶ Utiliser une régression GMZ régulière avec un estimateur de covariance Newey-West

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \gamma \mathbf{x}_t + \eta_t$$

$$H_0 : \beta_1 = \beta_2 = \gamma = 0, H_1 : \beta_1 \neq 0 \cup \beta_2 \neq 0 \cup \gamma_j \neq 0 \exists j$$

- ▶ Modéliser explicitement le  $MA(h-1)$  et utiliser un estimateur de covariance standard

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \gamma \mathbf{x}_t + \eta_t + \sum_{i=1}^{h-1} \theta_i \eta_{t-i}$$

Note : Nulle est la même ; n'impose pas de restriction sur  $\theta$

## Exemple : VAR de Politique Monétaire

- ▶ Prévisions produites de manière itérative pour 1 à 8 trimestres à l'avance
- ▶ Benchmark (*bm*) de marche aléatoire (*Fixed Forecast*) ou de moyenne constante
- ▶ AR et VAR sélectionnent la longueur de décalage en utilisant BIC
- ▶ Forcer la réversion à la moyenne dans l'échantillon en utilisant un estimateur à 2 étapes
  1. Estimer la moyenne de l'échantillon, et soustraire pour produire  $\tilde{\mathbf{y}}_t = \mathbf{y}_t - \hat{\boldsymbol{\mu}}$
  2. Estimer VAR *sans* constante

$$\tilde{\mathbf{y}}_t = \boldsymbol{\Phi}_1 \tilde{\mathbf{y}}_{t-1} + \dots + \boldsymbol{\Phi}_p \tilde{\mathbf{y}}_{t-p} + \boldsymbol{\varepsilon}_t$$

3. Prédiction puis ajouter la moyenne dans l'échantillon

$$\mathbf{E}_t [\tilde{\mathbf{y}}_{t+h}] + \hat{\boldsymbol{\mu}}$$

- ▶ Évaluation basée sur MSE relatif :

$$\text{Rel. MSE} = \frac{\text{MSE}}{\text{MSE}_{bm}}, \text{MSE} = 1/(T - h - R) \sum_{t=R}^{T-h} (y_{t+h} - \hat{y}_{t+h|t})^2$$

## Exemple : VAR de Politique Monétaire

Horizon	Série	VAR		AR	
		Restreinte	Non restreinte	Restreinte	Non restreinte
1	Chômage	0.522	0.520	0.507	0.507
	Taux des fonds fédéraux	0.887	0.903	0.923	0.933
	Inflation	0.869	0.868	0.839	0.840
2	Chômage	0.716	0.710	0.717	0.718
	Taux des fonds fédéraux	0.923	0.943	1.112	1.130
	Inflation	1.082	1.081	1.031	1.030
4	Chômage	0.872	0.861	0.937	0.940
	Taux des fonds fédéraux	0.952	0.976	1.082	1.109
	Inflation	1.000	0.999	0.998	0.998
8	Chômage	0.820	0.806	0.973	0.979
	Taux des fonds fédéraux	0.974	1.007	1.062	1.110
	Inflation	1.001	1.000	0.998	0.997

## Performances de prévision VAR vs AR

1. Le taux de chômage,
2. Le taux des Fed Funds,
3. L'inflation.

Les prévisions sont réalisées à partir de **50 % de l'échantillon disponible**, et les paramètres du modèle sont réestimés à chaque itération.

La précision des prévisions est mesurée par la **Mean Squared Error (MSE) hors échantillon**, comparée à celle d'un **modèle de référence** (random walk pour le Fed Funds rate et moyenne historique pour les autres variables).

Deux variantes du VAR sont testées :

- ▶ **Modèle restreint** : Forcé à converger vers la moyenne historique.
- ▶ **Modèle non restreint** : Estime librement les paramètres, y compris l'intercept.

Le VAR intègre des interactions entre plusieurs séries temporelles, offrant des prévisions plus précises que les modèles univariés dans 7 cas sur 12. Lorsqu'il n'est pas optimal, il reste compétitif.

En **politique monétaire** et plus généralement en macroéconomie, permet d'anticiper l'impact des chocs et améliorer la prise de décision.

## Estimation et Identification

- ▶ Identification univariée : Box-Jenkins
  - ▶ Utiliser ACF et PACF pour déterminer l'ordre de décalage AR et MA
  - ▶ Examiner les résidus
  - ▶ Principe de parcimonie
- ▶ L'autocorrélation d'un processus scalaire est définie

$$\rho_s = \frac{\gamma_s}{\gamma_0}$$

où  $\gamma_s$  est la  $s^{\text{ème}}$  autocovariance

- ▶ Coefficient de régression :

$$y_t = \mu + \rho_s y_{t-s} + \varepsilon_t$$

- ▶ Autocorrélation partielle  $\psi_s$ 
  - ▶ Interprétation de régression de la  $s^{\text{ème}}$  autocorrélation partielle :
$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{s-1} y_{t-s+1} + \psi_s y_{t-s} + \varepsilon_t$$
  - ▶  $\psi$  est la  $s^{\text{ème}}$  autocorrélation partielle

### (Fonction Cross Correlation) et (Fonction Cross Correlation Partielle)

- ▶ Équivalents multivariés
  - ▶ ACF et PACF ont les mêmes définitions de régression
  - ▶ Fonction de corrélation croisée

$$\rho_{xy,s} = \frac{E[(x_t - \mu_x)(y_{t-s} - \mu_y)]}{\sqrt{V[x_t] V[y_t]}}$$
$$\rho_{yx,s} = \frac{E[(y_t - \mu_y)(x_{t-s} - \mu_x)]}{\sqrt{V[x_t] V[y_t]}}$$

- ▶ Généralement différent
- ▶ Fonction de corrélation partielle croisée  $\psi_{xy,s}$

$$x_t = \phi_0 + \phi_{x1}x_{t-1} + \dots + \phi_{xs-1}x_{t-(s-1)} \\ + \phi_{y1}y_{t-1} + \dots + \phi_{ys-1}y_{t-(s-1)} + \varphi_{xy,s}y_{t-s} + \varepsilon_{x,t}$$

- ▶ Peut aider à identifier l'ordre VAR
- ▶ Problème plus profond : trop nombreux et trop compliqués
- ▶ Solution simple : Sélection de modèle



## Interprétation des CCF et PCCF

- $y$  a une dynamique (Hétérogène) HAR et se “déverse” dans  $x$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0.9 \\ .0 & 0.47 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{i=2}^5 \begin{bmatrix} 0 & 0 \\ 0 & 0.06 \end{bmatrix} \begin{bmatrix} x_{t-i} \\ y_{t-i} \end{bmatrix} \\ + \sum_{j=6}^{22} \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix} \begin{bmatrix} x_{t-j} \\ y_{t-j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

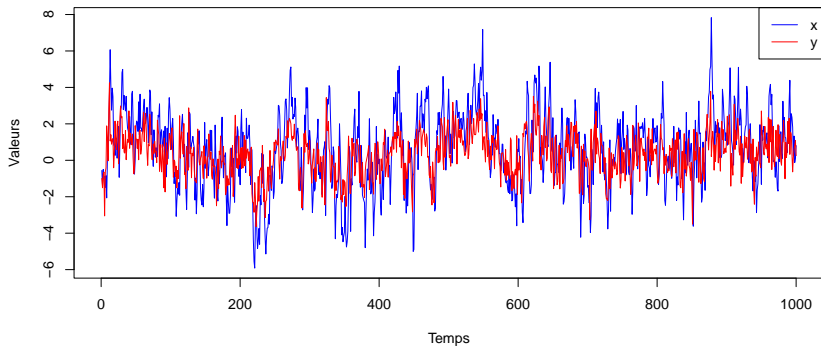
détaillée par

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \underbrace{\begin{bmatrix} 0.5 & 0.9 \\ 0 & 0.47 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix}}_{\text{Dépendance immédiate sur } t-1} + \underbrace{\sum_{i=2}^5 \begin{bmatrix} 0 & 0 \\ 0 & 0.06 \end{bmatrix} \begin{bmatrix} x_{t-i} \\ y_{t-i} \end{bmatrix}}_{\text{Effet de mémoire moyen terme (lags 2-5) sur } y_t} \\ + \underbrace{\sum_{j=6}^{22} \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix} \begin{bmatrix} x_{t-j} \\ y_{t-j} \end{bmatrix}}_{\text{Effet de mémoire long terme (lags 6-22) sur } y_t} + \underbrace{\begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}}_{\text{Innovations}}$$

## Graph des séries

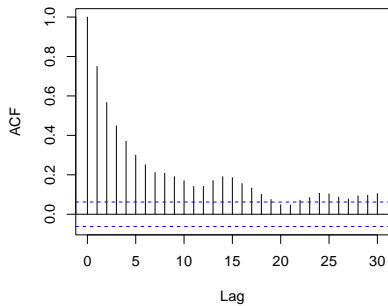
### ► Données simulées

**Simulation du modèle VAR**

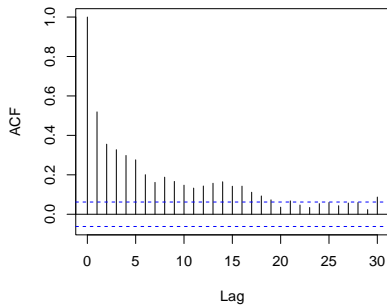


## ACF et CCFS

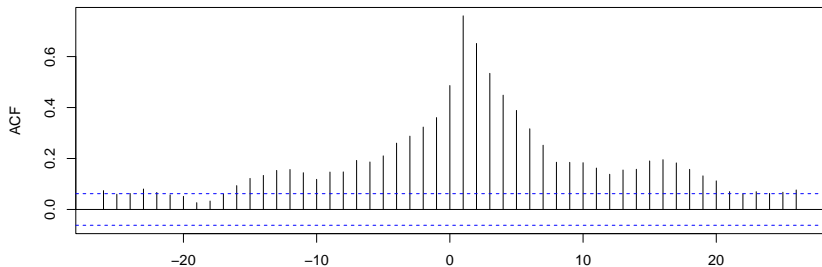
ACF de  $x_t$



ACF de  $y_t$

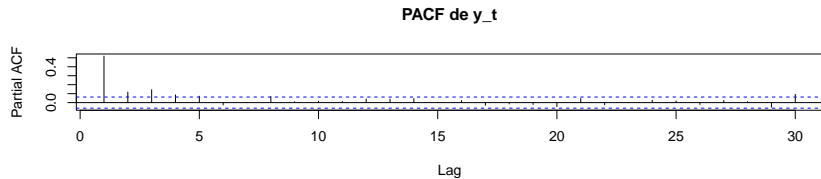
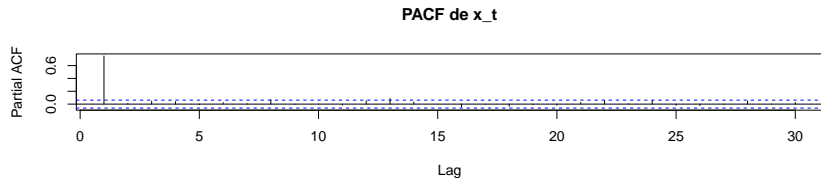


CCF entre  $x_t$  et  $y_t$



## Fonctions PACF

### PACF



## Sélection de modèle

- ▶ Étape 1 : Choisir la longueur de décalage maximale
  - ▶ Critères d'information

$$\begin{array}{ll}\text{AIC:} & \ln |\Sigma(P)| + k^2 P \frac{2}{T} \\ \text{Hannan-Quinn IC (HQIC):} & \ln |\Sigma(P)| + k^2 P \frac{\ln \ln T}{T} \\ \text{SIC:} & \ln |\Sigma(P)| + k^2 P \frac{\ln T}{T}\end{array}$$

- ▶  $\Sigma(P)$  est la covariance des résidus en utilisant  $P$  décalages
  - ▶  $|\cdot|$  est le déterminant
- ▶ Basé sur l'hypothèse de test
  - ▶ Général à Spécifique
  - ▶ Spécifique à Général
- ▶ Rapport de vraisemblance

$$\left( T - P_2 k^2 \right) \left( \ln |\Sigma(P_1)| - \ln |\Sigma(P_2)| \right) \overset{A}{\sim} \chi^2_{(P_2 - P_1)k^2}$$

## Sélection de la longueur de décalage dans le VAR de la Politique Monétaire

- Décalage maximal : 12 (1 an)

Longueur de décalage	AIC	HQIC	BIC	LR	P-val
0	4.014	3.762	3.605	925	0.000
1	0.279	0.079	0.000 <sup>▽Δ</sup>	39.6	0.000
2	0.190	0.042	0.041	40.9	0.000
3	0.096	0.000 <sup>▽</sup>	0.076	29.0	0.001
4	0.050 <sup>▽</sup>	0.007	0.160	7.34	0.602 <sup>▽</sup>
5	0.094	0.103	0.333	29.5	0.001
6	0.047	0.108	0.415	13.2	0.155
7	0.067	0.180	0.564	32.4	0.000
8	0.007	0.172 <sup>Δ</sup>	0.634	19.8	0.019
9	0.000 <sup>Δ</sup>	0.217	0.756	7.68	0.566 <sup>Δ</sup>
10	0.042	0.312	0.928	13.5	0.141
11	0.061	0.382	1.076	13.5	0.141
12	0.079	0.453	1.224	-	-

# Causalité de Granger

- ▶ Premier concept fondamentalement nouveau
- ▶ Examine si les décalages d'une variable sont utiles pour prédire une autre

## Définition (Causalité de Granger)

Une variable aléatoire scalaire  $\{x_t\}$  est dite **ne pas** causer au sens de Granger  $\{y_t\}$  si  $E[y_t \mid x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, \dots] = E[y_t \mid y_{t-1}, y_{t-2}, \dots]$ . C'est-à-dire,  $\{x_t\}$  ne cause pas au sens de Granger si la prévision de  $y_t$  est la même que l'on conditionne ou non sur les valeurs passées de  $x_t$ .

## Causalité de Granger

- ▶ Se traduit directement en une restriction dans un VAR
- ▶ Non restreint

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- ▶ Restreint de sorte que  $x_t$  ne GC pas  $y_t$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$x_t = \phi_{01} + \phi_{11}x_{t-1} + \phi_{12}y_{t-1} + \varepsilon_{1,t}$$

$$y_t = \phi_{02} + \phi_{22}y_{t-1} + \varepsilon_{2,t} \Leftarrow \text{Pas de } x_t!$$



## Plus de causalité de Granger

- Dans le modèle de décalage  $P$

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t$$

l'hypothèse nulle est

$$H_0 : \phi_{ij,1} = \phi_{ij,2} = \dots = \phi_{ij,P} = 0$$

- Alternative est

$$H_0 : \phi_{ij,1} \neq 0 \text{ ou } \phi_{ij,2} \neq 0 \text{ ou } \dots \text{ ou } \phi_{ij,P} \neq 0$$

- Test du rapport de vraisemblance

$$(T - Pk^2) (\ln |\Sigma_r| - \ln |\Sigma_u|) \overset{A}{\sim} \chi_P^2$$

- $\Sigma_u$  est la covariance des erreurs à partir du modèle non restreint
- $\Sigma_r$  est la covariance des erreurs à partir du modèle restreint
- $T - Pk^2$  est le nombre d'observations moins le nombre de paramètres libres dans le modèle non restreint
  - Pourquoi  $\chi_P^2$  ?

## VAR de Politique Monétaire – Campbell

- ▶ Outil standard dans l'analyse de la politique monétaire
  - ▶ Taux de chômage (différencié)
  - ▶ Taux des fonds fédéraux
  - ▶ Taux d'inflation (différencié)

$$\begin{bmatrix} \Delta \text{UNEMP}_t \\ \text{FF}_t \\ \Delta \text{INF}_t \end{bmatrix} = \Phi_0 + \Phi_1 \begin{bmatrix} \Delta \text{UNEMP}_{t-1} \\ \text{FF}_{t-1} \\ \Delta \text{INF}_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix}$$

## Causalité de Granger dans le VAR de Campbell

- Utiliser le modèle avec 3 décalages (HQIC)
- $H_0 : \phi_{ij,1} = \phi_{ij,2} = \phi_{ij,3} = 0$
- $H_1 : \phi_{ij,1} \neq 0 \text{ ou } \phi_{ij,2} \neq 0 \text{ ou } \phi_{ij,3} \neq 0$
- $i$  représente la série affectée par les décalages de la série  $j$

	Taux des fonds fédéraux		Inflation		Chômage	
Exclusion	P-val	Stat	P-val	Stat	P-val	Stat
Taux des fonds fédéraux	-	-	0.001	13.068	0.014	8.560
Inflation	0.001	14.756	-	-	0.375	1.963
Chômage	0.000	19.586	0.775	0.509	-	-
Tous	0.000	33.139	0.000	18.630	0.005	10.472

# Fonctions de réponse impulsionnelle

- ▶ Deuxième concept fondamentalement nouveau
- ▶ La dynamique complexe d'un VAR rend l'interprétation directe des coefficients difficile
- ▶ La solution est d'examiner les réponses impulsionnelles
- ▶ La fonction de réponse impulsionnelle de  $y_i$  par rapport à un choc dans  $\varepsilon_j$ , pour tout  $j$  et  $i$ , est définie comme le changement dans  $y_{it+s}$ ,  $s \geq 0$  pour un choc unitaire dans  $\varepsilon_{jt}$ 
  - ▶ Difficile à déchiffrer
- ▶ Tant que  $\mathbf{y}_t$  est stationnaire en covariance, il doit avoir une représentation VMA,

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Xi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Xi}_2 \boldsymbol{\varepsilon}_{t-2} + \dots$$

- ▶  $\boldsymbol{\Xi}_j$  sont les réponses impulsionnelles !
- ▶ Pourquoi ?
  - ▶ Mesurent directement l'effet dans la période  $j$  de tout choc

## AP(P) et MA( $\infty$ )

- ▶ Tout AR( $P$ ) stationnaire

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_P y_{t-P} + \varepsilon_t$$

peut être représenté comme un MA( $\infty$ )

$$y_t = \phi_0 / (1 - \phi_1 - \phi_2 - \dots - \phi_P) + \varepsilon_t + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i}$$

- ▶ AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

devient

$$y_t = \phi_0 / (1 - \phi_1) + \varepsilon_t + \sum_{i=1}^{\infty} \phi_1^i \varepsilon_{t-i}$$

- ▶ Les VAR( $P$ ) stationnaires ont la même relation avec VMA( $\infty$ )

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t$$

$$\mathbf{y}_t = \boldsymbol{\mu} + \varepsilon_t + \Xi_1 \varepsilon_{t-1} + \Xi_2 \varepsilon_{t-2} + \dots$$

## Résolution IR

- Facile dans VAR(1)

$$y_t = (\mathbf{I}_K - \Phi_1)^{-1} \Phi_0 + \varepsilon_t + \Phi_1 \varepsilon_{t-1} + \Phi_1^2 \varepsilon_{t-2} + \dots$$

- $\Xi_j = \Phi_1^j$
- Dans le VAR(P) général,

$$\Xi_j = \Phi_1 \Xi_{j-1} + \Phi_2 \Xi_{j-2} + \dots + \Phi_P \Xi_{j-P}$$

où  $\Xi_0 = \mathbf{I}_k$  et  $\Xi_m = 0$  pour  $m < 0$ .

- Dans un VAR(2),

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \varepsilon_t$$
$$-\Xi_0 = \mathbf{I}_k, \Xi_1 = \Phi_1, \Xi_2 = \Phi_1^2 + \Phi_2, \text{ et } \Xi_3 = \Phi_1^3 + \Phi_1 \Phi_2 + \Phi_2 \Phi_1$$

- Les intervalles de confiance sont également assez compliqués

## Considérations pour les chocs

- ▶ VAR bivariée simple d'ordre 1

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- ▶ La manière dont vous choquez importe
- ▶ Dépend de la corrélation entre  $\varepsilon_{1,t}$  et  $\varepsilon_{2,t}$
- ▶ 3 méthodes
  - ▶ Ignorer la corrélation et choquer simplement  $\varepsilon_{j,t}$  avec un choc d'écart-type unitaire
  - ▶ Utiliser Cholesky pour factoriser  $\Sigma$  et utiliser  $\Sigma^{1/2} \mathbf{e}_j$  où  $\mathbf{e}_j$  est un vecteur de zéros avec 1 dans la  $j^{\text{ème}}$  position

$$\Sigma = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix} \quad \Sigma_C^{1/2} = \begin{bmatrix} 1 & 0 \\ .5 & .866 \end{bmatrix}$$

- ▶ L'ordre des variables importe
- ▶ “Réponse impulsionnelle généralisée” qui utilise une méthode de projection

## Exemple des différents chocs

- Définir la covariance des erreurs

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{bmatrix}$$

- Standardisé

$$\begin{bmatrix} \sigma_x \\ 0 \end{bmatrix} \text{ et } \begin{bmatrix} 0 \\ \sigma_y \end{bmatrix}$$

- Cholesky

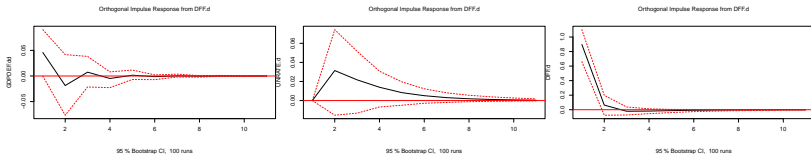
$$\Sigma_C^{1/2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_x & 0 \\ \sigma_y \rho & \sigma_y \sqrt{1 - \rho^2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma_y \sqrt{1 - \rho^2} \end{bmatrix}$$
$$\begin{bmatrix} \sigma_x & 0 \\ \sigma_y \rho & \sigma_y \sqrt{1 - \rho^2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_x \\ \sigma_y \rho \end{bmatrix}, \text{ autre est } \begin{bmatrix} 0 \\ \sigma_y \sqrt{1 - \rho^2} \end{bmatrix}$$



## Réponses impulsionnelles

- ▶ Taux des fonds fédéraux ordonné en premier
- ▶ Réponse à un choc des fonds fédéraux
- ▶ Factorisation de Cholesky

```
irf_var_FF_INF <- irf(PolMon, impulse = "DFF.d", response = "GDPDEF.dd")
irf_var_FF_UNRATE <- irf(PolMon, impulse = "DFF.d", response = "UNRATE.d")
irf_var_FF_FF <- irf(PolMon, impulse = "DFF.d", response = "DFF.d")
par(mfrow=c(1,3))
plot(irf_var_FF_INF)
plot(irf_var_FF_UNRATE)
plot(irf_var_FF_FF)
par(mfrow=c(1,1))
```



## Cointégration

- ▶ La cointégration est la version VAR des racines unitaires
- ▶ Établit des relations à long terme entre deux variables à racine unitaire
  - ▶ La consommation a une racine unitaire, le revenu a une racine unitaire
  - ▶ Consommation - Revenu : ????

## Définition (Intégré d'Ordre 1)

Une variable  $y_t$  est intégrée d'ordre 1, ou  $I(1)$ , si  $y_t$  est non stationnaire et  $\Delta y_t = y_t - y_{t-1}$  est stationnaire.

## Cointégration

### Définition (Cointégration Bivariée)

Si  $x_t$  et  $y_t$  sont cointégrés si les deux sont  $I(1)$  et il existe un vecteur  $\beta$  avec les deux éléments non nuls tels que

$$\beta_1 x_t - \beta_2 y_t \sim I(0)$$

- ▶ Lien fort entre  $x_t$  et  $y_t$
- ▶ Les deux sont des marches aléatoires mais la différence est à retour moyen
- ▶ Retour moyen à la tendance (tendance stochastique)

À quoi ressemble la cointégration ?

$$\mathbf{y}_t = \Phi_{ij} \mathbf{y}_{t-1} + \varepsilon_t$$

$$\Phi_{11} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}, \quad \lambda_i = 1, 0.6$$

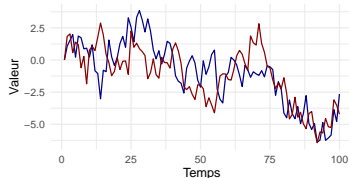
$$\Phi_{21} = \begin{bmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{bmatrix}, \quad \lambda_i = 0.9, 0.5$$

$$\Phi_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \lambda_i = 1, 1$$

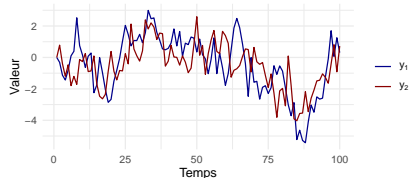
$$\Phi_{22} = \begin{bmatrix} -0.3 & 0.3 \\ 0.1 & -0.2 \end{bmatrix}, \quad \lambda_i = -0.43, -0.06$$

# Persistence, Anti-persistence et Cointégration

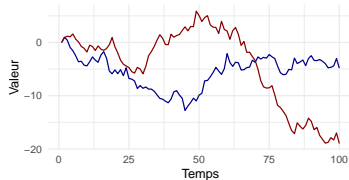
Cointegration ( $\Phi_{11}$ )



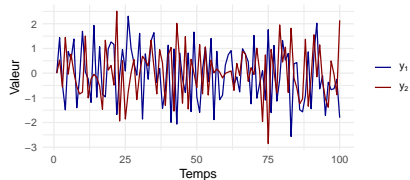
Persistant, Stationnaire ( $\Phi_{21}$ )



Racines unitaires indépendantes ( $\Phi_{12}$ )



Antipersistant, Stationnaire avec ( $\Phi_{22}$ )



## Comment savons-nous quand un VAR est cointégré ?

- La condition des valeurs propres détermine si un VAR(1) est cointégré

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- Cointégré si une seule valeur propre est unitaire.
- Si tous inférieurs à 1 : ??
- Si les deux 1 : deux racines unitaires indépendantes

$$\Phi_{11} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \quad \Phi_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\lambda_i = 1, 0.6$$

$$\lambda_i = 1, 1$$

$$\Phi_{21} = \begin{bmatrix} .7 & .2 \\ .2 & .7 \end{bmatrix} \quad \Phi_{22} = \begin{bmatrix} -.3 & .3 \\ .1 & -.2 \end{bmatrix}$$

$$\lambda_i = 0.9, 0.5$$

$$\lambda_i = -0.43, -0.06$$

# Modèles de Correction d'Erreur

- ▶ Point majeur de la cointégration
  - ▶ Cointégré  $\Leftrightarrow$  Modèle de correction d'erreur
- ▶ Qu'est-ce qu'un modèle de correction d'erreur ?
  - ▶ VAR cointégré :

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- ▶ Modèle de correction d'erreur :

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 & .2 \\ .2 & -.2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- ▶ Forme normalisée

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 \\ .2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- ▶  $\begin{bmatrix} 1 & -1 \end{bmatrix}$  est le vecteur de cointégration
- ▶  $\begin{bmatrix} -.2 & .2 \end{bmatrix}'$  mesure la vitesse d'ajustement

## De VAR à VECM

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

Soustraire  $[y_{t-1} \ x_{t-1}]'$  des deux côtés

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} - \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} - \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \left( \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 & .2 \\ .2 & -.2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 \\ .2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$



## Vecteurs de cointégration

- ▶ La relation de cointégration peut toujours être décomposée

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \varepsilon_t$$

$$\boldsymbol{\pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$$

- ▶  $\boldsymbol{\alpha}$  mesure la vitesse de convergence
- ▶  $\boldsymbol{\beta}$  contiennent les vecteurs de cointégration
- ▶ Le nombre de vecteurs de cointégration est  $\text{rank}(\boldsymbol{\alpha} \boldsymbol{\beta}')$

$$\boldsymbol{\alpha} \boldsymbol{\beta}' = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

- ▶ Combien ?

## Détermination des vecteurs de cointégration

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \varepsilon_t$$
$$\boldsymbol{\pi} = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

- Mettre  $\boldsymbol{\pi}$  sous forme échelonnée par ligne

$$\text{Forme Échelonnée par Ligne} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.3 \\ 0 & 0 & 0 \end{bmatrix}$$

- Rappel  $\boldsymbol{\pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -0.3 \end{bmatrix} \quad \boldsymbol{\alpha} = \begin{bmatrix} .3 & .2 \\ .2 & .5 \\ -.3 & -.3 \end{bmatrix}$$

## Résolution des vecteurs de cointégration

$$\alpha\beta' = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

$$\text{Forme Échelonnée par Ligne} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \beta_1 & \beta_2 \end{bmatrix}$$

et  $\alpha$  a 6 paramètres inconnus.  $\alpha\beta'$  peut être combiné pour produire

$$\pi = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{11}\beta_1 + \alpha_{12}\beta_2 \\ \alpha_{21} & \alpha_{22} & \alpha_{21}\beta_1 + \alpha_{22}\beta_2 \\ \alpha_{31} & \alpha_{32} & \alpha_{31}\beta_1 + \alpha_{32}\beta_2 \end{bmatrix}$$

## Tester la cointégration

- ▶ Deux tests pour la cointégration
  - ▶ Engle-Granger
  - ▶ Johansen
- ▶ Nous allons nous concentrer sur Engle-Granger
  - ▶ Simple et intuitif
  - ▶ Seulement applicable avec 1 relation de cointégration
- ▶ Tester la propriété clé de la cointégration : la différence est  $I(0)$
- ▶ La plupart du travail est une simple OLS

$$y_t = \delta_0 + \beta x_t + \varepsilon_t$$

- ▶ Reste du travail est de tester  $\hat{\varepsilon}_t$  pour une racine unitaire
- ▶ Johansen teste les valeurs propres de  $\pi = \alpha\beta'$  directement.

## Algorithme (Test Engle-Granger)

1. Commencer par analyser  $x_t$  et  $y_t$  isolément. Les deux doivent être des racines unitaires pour envisager la cointégration.
2. Estimer la relation à long terme

$$y_t = \delta_0 + \beta x_t + \varepsilon_t$$

et tester  $H_0 : \gamma = 0$  contre  $H_0 : \gamma < 0$  dans la régression ADF

$$\Delta \hat{\varepsilon}_t = \gamma \hat{\varepsilon}_{t-1} + \delta_1 \Delta \hat{\varepsilon}_{t-1} + \dots + \delta_p \Delta \hat{\varepsilon}_{t-p} + \eta_t$$

3. En utilisant les paramètres estimés, spécifier et estimer la forme de correction d'erreur de la relation,

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} \pi_{01} \\ \pi_{02} \end{bmatrix} + \begin{bmatrix} \alpha_1 \hat{\varepsilon}_t \\ \alpha_2 \hat{\varepsilon}_t \end{bmatrix} + \pi_1 \begin{bmatrix} \Delta x_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \dots + \pi_p \begin{bmatrix} \Delta x_{t-p} \\ \Delta y_{t-p} \end{bmatrix} + \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \end{bmatrix}$$

4. Évaluer le modèle

## Considérations Engle-Granger

- ▶ Termes déterministes

- ▶ Pas de termes déterministes : seulement dans des circonstances spéciales

$$y_t = \beta x_t + \varepsilon_t$$

- ▶ Constante : cas standard

$$y_t = \delta_0 + \beta x_t + \varepsilon_t$$

- ▶ Constante et tendance temporelle : permettre différents taux de croissance/tendances temporelles dans les variables

$$y_t = \delta_0 + \delta_1 t + \beta x_t + \varepsilon_t$$

## Valeurs Critiques

- ▶ Les valeurs critiques dépendent des déterministes dans la régression CI
  - ▶ Les modèles avec plus de déterministes ont des valeurs critiques plus basses (plus négatives)
- ▶ Les valeurs critiques dépendent du nombre de variables  $I(1)$  du côté droit
  - ▶ Les modèles plus grands ont des valeurs critiques plus basses

## Exemple : cay

- ▶ La relation consommation-richesse agrégée a été une série cointégrée intéressante dans la littérature financière récente
- ▶ A relancé le CCAPM
- ▶ Trois composantes :
  - ▶ Consommation ( $c$ )
  - ▶ Richesse en actifs ( $a$ )
  - ▶ Revenu du travail (richesse humaine) ( $y$ )
- ▶ Écart par rapport à la relation à long terme lié au rendement attendu
- ▶ Relation de cointégration :  $c_t + 0.643 - 0.249a_t - 0.785y_t$

Tests de racine unitaire			
Série	T-stat	P-val	Décalages ADF
$c$	-1.198	0.674	5
$a$	-0.205	0.938	3
$y$	-2.302	0.171	0
$\hat{\varepsilon}_t^c$	-2.706	0.383	1
$\hat{\varepsilon}_t^a$	-2.573	0.455	0
$\hat{\varepsilon}_t^y$	-2.679	0.398	1

```
# Télécharger les données
consommation <- fredr(series_id = "PCE", observation_start = as.Date("1955-01-01"), observation_end = as.Date("2015-01-01"), frequency = "quarterly", file_format = "csv")
richesse <- fredr(series_id = "TNWBSHNO", observation_start = as.Date("1955-01-01"), observation_end = as.Date("2015-01-01"), frequency = "quarterly", file_format = "csv")
revenu <- fredr(series_id = "PI", observation_start = as.Date("1955-01-01"), observation_end = as.Date("2015-01-01"), frequency = "quarterly", file_format = "csv")
```

```
# Fusionner les données
```

```
donnees <- consommation %>%
  select(date, valeur_c = value) %>%
  left_join(richesse %>% select(date, valeur_a = value), by = "date") %>%
  left_join(revenu %>% select(date, valeur_y = value), by = "date")
```

```
# Nettoyer les données
```

```
donnees <- na.omit(donnees)
```

```
donnees <- donnees %>%
```

```
  mutate(
    log_c = log(valeur_c),
    log_a = log(valeur_a),
    log_y = log(valeur_y)
  )
```

```
# Test de stationnarité
```

```
adf.test(donnees$log_c)
```

```
## Augmented Dickey-Fuller Test
```

```
## alternative: stationary
```

```
##
```

```
## Type 1: no drift no trend
```

```
##      lag    ADF p.value
```

```
## [1,]  0 22.22   0.99
```



## Modèle Vectoriel de Correction d'Erreur

- ▶ VECM estimé en utilisant les résidus de la régression de cointégration

$$\begin{bmatrix} \Delta c_t \\ \Delta a_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} 0.003 \\ (0.000) \\ 0.004 \\ (0.014) \\ 0.003 \\ (0.000) \end{bmatrix} + \begin{bmatrix} -0.000 \\ (0.281) \\ 0.002 \\ (0.037) \\ 0.000 \\ (0.515) \end{bmatrix} \hat{\varepsilon}_{t-1} + \begin{bmatrix} 0.192 & 0.102 & 0.139 \\ 0.005 & (0.000) & (0.004) \\ 0.282 & 0.220 & -0.149 \\ (0.116) & (0.006) & (0.414) \\ 0.369 & 0.061 & -0.139 \\ (0.000) & (0.088) & (0.140) \end{bmatrix} \begin{bmatrix} \Delta c_{t-1} \\ \Delta a_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \eta_t$$

- ▶ Valeurs p entre parenthèses
- ▶ Estimation de la relation de cointégration n'a aucun effet sur les erreurs standard
  - ▶ Converge rapidement ( $T$ )
  - ▶ Les paramètres VECM convergent à la racine  $\sqrt{T}$

## Régression Fallacieuse et Équilibre

- ▶ La prudence est nécessaire lorsqu'on travaille avec des données  $I(1)$ 
  - ▶  $I(0)$  sur  $I(0)$  : Le cas habituel. Les arguments asymptotiques standards s'appliquent.
  - ▶  $I(1)$  sur  $I(0)$  : Cette régression est déséquilibrée.
  - ▶  $I(1)$  sur  $I(1)$  : Cointégration ou régression fallacieuse.
  - ▶  $I(0)$  sur  $I(1)$  : Cette régression est déséquilibrée.
- ▶ La régression fallacieuse peut conduire à de grandes statistiques  $t$  lorsque les séries sont indépendantes.
  - ▶ Deux processus  $I(1)$  non liés,  $x_t$  et  $y_t$  :

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

$$y_t = y_{t-1} + \eta_t$$

- ▶ Lorsque  $T = 50$ , environ 80% des statistiques  $t$  sont significatives
- ▶ Toujours vérifier pour  $I(1)$  lors de l'utilisation de données de séries temporelles
- ▶ Si les deux sont  $I(1)$ , assurez-vous qu'ils sont cointégrés.