# Adaptive Estimation of VAR with Time-Varying Variance : Application to Testing Linear Causality in Mean and VAR Order

V. Patilea (CREST-Ensai & IRMAR, Rennes, France)

*joint work with*

H. Raïssi (IRMAR-INSA)

JMS 2012, INSEE

# Outline

# Outline

# Outline

1. The model
2. Least squares parameter estimation
3. Adaptive Least Squares parameter estimation

# Outline

# Outline

# Outline

- Observations $X_{-p+1}, \ldots, X_0, X_1, \ldots, X_T \in \mathbb{R}^d$

- Observations $X_{-p+1}, \ldots, X_0, X_1, \ldots, X_T \in \mathbb{R}^d$
- Multivariate time-series model: linear VAR

$$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-p} + u_t$$
$$u_t = H_t \epsilon_t,$$

- Observations $X_{-p+1}, \dots, X_0, X_1, \dots, X_T \in \mathbb{R}^d$
- Multivariate time-series model: linear VAR

$$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-p} + u_t$$
$$u_t = H_t \epsilon_t,$$

- Stability condition:

$$\det A(z) \neq 0 \quad \text{for all } |z| \leq 1 \text{ where } A(z) = I_d - \sum_{i=1}^{p} A_i z^i$$

- Let $\mathcal{F}_t = \sigma\{\epsilon_s : s \leq t\}$.

### Assumption A1:

- The $d \times d$ matrices $H_t$ are invertible and $H_t = G(t/T)$, $1 \leq t \leq T$.

- Let $\mathcal{F}_t = \sigma\{\epsilon_s : s \leq t\}$.

## Assumption A1:

- The $d \times d$ matrices $H_t$ are invertible and $H_t = G(t/T)$, $1 \leq t \leq T$.
- The components of $G(r) := \{g_{kl}(r)\}$ are *deterministic* functions on the interval $(0, 1]$ and

- Let $\mathcal{F}_t = \sigma\{\epsilon_s : s \leq t\}$.

**Assumption A1:**

- The $d \times d$ matrices $H_t$ are invertible and $H_t = G(t/T)$, $1 \leq t \leq T$.
- The components of $G(r) := \{g_{kl}(r)\}$ are *deterministic* functions on the interval $(0, 1]$ and
- The process $(\epsilon_t)$ is $\alpha$-mixing and

- Let $\mathcal{F}_t = \sigma\{\epsilon_s : s \leq t\}$.

**Assumption A1:**

- The $d \times d$ matrices $H_t$ are invertible and $H_t = G(t/T)$, $1 \leq t \leq T$.
- The components of $G(r) := \{g_{kl}(r)\}$ are *deterministic* functions on the interval $(0, 1]$ and
  - $\sup_{r\in(0,1]} |g_{kl}(r)| < \infty$,
  - $g_{kl}(\cdot)$ are piecewise Lipschitz continuous on $(0, 1]$,
  - $\Sigma(\cdot) = G(\cdot)G(\cdot)' \gg 0$ for all $r$.
- The process $(\epsilon_t)$ is $\alpha$-mixing and
  - $E(\epsilon_t \mid \mathcal{F}_{t-1}) = 0$,
  - $E(\epsilon_t\epsilon_t' \mid \mathcal{F}_{t-1}) = I_d$,
  - the $d$ components $\epsilon_{kt}$ of $(\epsilon_t)$ satisfy $\sup_t \parallel \epsilon_{kt} \parallel_{4\mu} < \infty$, $\mu > 1$.

**Remarks:**

- Unconditional non-stationary volatility (time-varying variance)
- Weak regularity conditions on the time-varying variance
- Multivariate GARCH structure cannot be taken into account

# Outline

- The OLS estimator

$$\hat{\theta}_{OLS} = \hat{\Sigma}_{\tilde{X}}^{-1} \text{vec}\left(\hat{\Sigma}_X\right),$$

where

$$\hat{\Sigma}_{\tilde{X}} = T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1} \tilde{X}'_{t-1} \otimes I_d \quad \text{and} \quad \hat{\Sigma}_X = T^{-1} \sum_{t=1}^{T} X_t \tilde{X}'_{t-1}$$

and $\tilde{X}_{t-1} = (X'_{t-1}, \ldots, X'_{t-p})'$.

- Let $\Sigma_t := H_t H_t'$ (unconditional time-varying variance)

- Let $\Sigma_t := H_t H_t'$ (unconditional time-varying variance)
- The Generalized Least Squares (GLS) estimator

$$\hat{\theta}_{GLS} = \hat{\Sigma}_{\underline{\tilde{X}}}^{-1} \text{vec}\left(\hat{\Sigma}_{\underline{X}}\right),$$

with

$$\hat{\Sigma}_{\underline{\tilde{X}}} = T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1} \tilde{X}_{t-1}' \otimes \Sigma_t^{-1} \quad \text{and} \quad \hat{\Sigma}_{\underline{X}} = T^{-1} \sum_{t=1}^{T} \Sigma_t^{-1} X_t \tilde{X}_{t-1}'.$$

- If the volatility matrix $\Sigma_t$ is constant in time, $\hat{\theta}_{GLS} = \hat{\theta}_{OLS}$.

- If the volatility matrix $\Sigma_t$ is constant in time, $\hat{\theta}_{GLS} = \hat{\theta}_{OLS}$.
- The GLS estimator is in general infeasible.

# Outline

## Theorem

*If Assumption* **A1** *holds true, then:*

1.
$$T^{\frac{1}{2}}(\hat{\theta}_{OLS} - \theta_0) \Rightarrow \mathcal{N}(0, \Lambda_3^{-1}\Lambda_2\Lambda_3^{-1}),$$

*where*
$$\Lambda_2 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p\times p} \otimes \Sigma(r))\tilde{\psi}_i' \right\} \otimes \Sigma(r)dr$$

*and*
$$\Lambda_3 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p\times p} \otimes \Sigma(r))\tilde{\psi}_i' \right\} \otimes I_d \, dr$$

*are positive definite;*

## Theorem (continued)

2.
$$T^{\frac{1}{2}}(\hat{\theta}_{GLS} - \theta_0) \Rightarrow \mathcal{N}(0, \Lambda_1^{-1}),$$

*where*

$$\Lambda_1 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i (\mathbf{1}_{p\times p} \otimes \Sigma(r)) \tilde{\psi}_i' \right\} \otimes \Sigma(r)^{-1} dr$$

*is positive definite;*

## Theorem (continued)

2.
$$T^{\frac{1}{2}}(\hat{\theta}_{GLS} - \theta_0) \Rightarrow \mathcal{N}(0, \Lambda_1^{-1}),$$

*where*

$$\Lambda_1 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i (\mathbf{1}_{p\times p} \otimes \Sigma(r)) \tilde{\psi}_i' \right\} \otimes \Sigma(r)^{-1} dr$$

*is positive definite;*

3. *The asymptotic variance of $\hat{\theta}_{GLS}$ is smaller than the asymptotic variance of $\hat{\theta}_{OLS}$, that is*

$$\Lambda_3^{-1}\Lambda_2\Lambda_3^{-1} - \Lambda_1^{-1} \gg 0.$$

# Remarks:

- In the homoscedastic (time-constant variance) case, if $\Sigma(r) \equiv \Sigma_u$

$$\Lambda_1^{-1} = \Lambda_3^{-1}\Lambda_2\Lambda_3^{-1} = \{E[\tilde{X}_t\tilde{X}_t']\}^{-1} \otimes \Sigma_u, \tag{1}$$

# Remarks:

- In the homoscedastic (time-constant variance) case, if $\Sigma(r) \equiv \Sigma_u$

$$\Lambda_1^{-1} = \Lambda_3^{-1}\Lambda_2\Lambda_3^{-1} = \{E[\tilde{X}_t\tilde{X}_t']\}^{-1} \otimes \Sigma_u, \qquad (1)$$

- If $\Sigma(r) = \sigma^2(r)I_d$, the GLS asymp. variance does not depend on $\Sigma(r)$ (in particular if $d = 1$).

# Remarks:

- In the homoscedastic (time-constant variance) case, if $\Sigma(r) \equiv \Sigma_u$

$$\Lambda_1^{-1} = \Lambda_3^{-1}\Lambda_2\Lambda_3^{-1} = \{E[\tilde{X}_t\tilde{X}_t']\}^{-1} \otimes \Sigma_u, \tag{1}$$

- If $\Sigma(r) = \sigma^2(r)I_d$, the GLS asymp. variance does not depend on $\Sigma(r)$ (in particular if $d = 1$).

  However, in general this is not the case!

# Remarks:

- In the homoscedastic (time-constant variance) case, if $\Sigma(r) \equiv \Sigma_u$

$$\Lambda_1^{-1} = \Lambda_3^{-1}\Lambda_2\Lambda_3^{-1} = \{E[\tilde{X}_t\tilde{X}_t']\}^{-1} \otimes \Sigma_u, \tag{1}$$

- If $\Sigma(r) = \sigma^2(r)I_d$, the GLS asymp. variance does not depend on $\Sigma(r)$ (in particular if $d = 1$).

  However, in general this is not the case!

- In the case $d = 1$ we recover the results of Xu & Phillips (2008).

# Variance estimation for $\hat{\theta}_{OLS}$

## Proposition

*Under Assumption* **A1**, *if $\hat{u}_t$ denotes OLS residuals*

$$\hat{\Lambda}_2 := T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1}\tilde{X}_{t-1}' \otimes \hat{u}_t\hat{u}_t' = \Lambda_2 + o_p(1),$$

$$\hat{\Lambda}_3 := T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1}\tilde{X}_{t-1}' \otimes I_d = \Lambda_3 + o_p(1).$$

# Outline

- GLS estimator is generally infeasible

- GLS estimator is generally infeasible
- Idea: consider a nonparametric estimation of the volatility function

- GLS estimator is generally infeasible
- Idea: consider a nonparametric estimation of the volatility function
- Xu & Phillips (2008) proposed this approach in the case $d = 1$

- Define the symmetric matrix

$$\check{\Sigma}_t^0 = \sum_{i=1}^{T} w_{ti} \odot \hat{u}_i \hat{u}_i',$$

where

- Define the symmetric matrix

$$\check{\Sigma}_t^0 = \sum_{i=1}^{T} w_{ti} \odot \hat{u}_i \hat{u}_i',$$

where

  - $\hat{u}_i$'s are the OLS residuals

- Define the symmetric matrix

$$\check{\Sigma}_t^0 = \sum_{i=1}^{T} w_{ti} \odot \hat{u}_i \hat{u}_i',$$

where

- $\hat{u}_i$'s are the OLS residuals
- $\odot$ denotes the Hadamard (entrywise) product

- Define the symmetric matrix

$$\check{\Sigma}_t^0 = \sum_{i=1}^{T} w_{ti} \odot \hat{u}_i \hat{u}_i',$$

where

  - $\hat{u}_i$'s are the OLS residuals
  - $\odot$ denotes the Hadamard (entrywise) product
  - the $kl-$element, $k \leq l$, of the $d \times d$ weight matrix $w_{ti}$ is

$$w_{ti}(b_{kl}) = \left( \sum_{i=1}^{T} K_{ti}(b_{kl}) \right)^{-1} K_{ti}(b_{kl}),$$

with $b_{kl}$ the bandwidth and

$$K_{ti}(b_{kl}) = \begin{cases} K(\frac{t-i}{Tb_{kl}}) & \text{if} \quad t \neq i, \\ 0 & \text{if} \quad t = i. \end{cases}$$

- The kernel is a bounded density satisfying mild conditions

- The kernel is a bounded density satisfying mild conditions
- The bandwidths can be different from cell to cell

- The kernel is a bounded density satisfying mild conditions
- The bandwidths can be different from cell to cell
  - $b_{kl} \in \mathcal{B}_T$, $1 \leq k \leq l \leq d$, where $\mathcal{B}_T = [c_{min} b_T, c_{max} b_T]$ with $0 < c_{min} < c_{max} < \infty$
  - For some $\gamma > 0$, $b_T + 1/Tb_T^{2+\gamma} \to 0$ as $T \to \infty$.

# Bandwidth practical issues

- The bandwidths $b_{kl}$ can be chosen by CV, i.e. minimization of

$$\sum_{t=1}^{T} \| \check{\Sigma}_t - \hat{u}_t \hat{u}_t' \|^2,$$

w.r.t. all $b_{kl} \in \mathcal{B}_T$, $1 \leq k \leq l \leq d$, where $\| \cdot \|$ is the Frobenius norm.

# Bandwidth practical issues

- The bandwidths $b_{kl}$ can be chosen by CV, i.e. minimization of

$$\sum_{t=1}^{T} \parallel \check{\Sigma}_t - \hat{u}_t \hat{u}_t' \parallel^2,$$

w.r.t. all $b_{kl} \in \mathcal{B}_T$, $1 \leq k \leq l \leq d$, where $\| \cdot \|$ is the Frobenius norm.

- Theoretical results will be obtained uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$

# Bandwidth practical issues

- The bandwidths $b_{kl}$ can be chosen by CV, i.e. minimization of

$$\sum_{t=1}^{T} \| \check{\Sigma}_t - \hat{u}_t \hat{u}_t' \|^2,$$

w.r.t. all $b_{kl} \in \mathcal{B}_T$, $1 \leq k \leq l \leq d$, where $\| \cdot \|$ is the Frobenius norm.

- Theoretical results will be obtained uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ $\Rightarrow$ justification of the common cross-validation bandwidth rule

# The ALS estimator

- $$\hat{\theta}_{ALS} = \check{\Sigma}_{\underline{\tilde{X}}}^{-1} \text{vec}\left(\check{\Sigma}_{\underline{X}}\right),$$

# The ALS estimator

- $$\hat{\theta}_{ALS} = \check{\Sigma}_{\underline{\tilde{X}}}^{-1} \text{vec}\left(\check{\Sigma}_{\underline{X}}\right),$$

- $$\check{\Sigma}_{\underline{\tilde{X}}} = T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1} \tilde{X}_{t-1}' \otimes \check{\Sigma}_t^{-1} \quad \text{and} \quad \check{\Sigma}_{\underline{X}} = T^{-1} \sum_{t=1}^{T} \check{\Sigma}_t^{-1} X_t \tilde{X}_{t-1}'.$$

# The ALS estimator

- $$\hat{\theta}_{ALS} = \check{\Sigma}_{\underline{\tilde{X}}}^{-1} \text{vec}\left(\check{\Sigma}_{\underline{X}}\right),$$

- $$\check{\Sigma}_{\underline{\tilde{X}}} = T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1} \tilde{X}_{t-1}' \otimes \check{\Sigma}_t^{-1} \quad \text{and} \quad \check{\Sigma}_{\underline{X}} = T^{-1} \sum_{t=1}^{T} \check{\Sigma}_t^{-1} X_t \tilde{X}_{t-1}'.$$

- The ALS estimator $\hat{\theta}_{ALS}$ depends on the bandwidths!

# Asymptotic normality and variance estimation

## Theorem

*If*

- *a strengthened version of Assumption A1 holds true*
- *conditions on $K(\cdot)$ and $b_{kl}$ hold true*

*then uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ as $T \to \infty$*

$$\check{\Lambda}_1 = T^{-1} \sum_{t=1}^{T} \tilde{X}_{t-1} \tilde{X}'_{t-1} \otimes \check{\Sigma}_t^{-1} = \Lambda_1 + o_p(1),$$

*and*

$$\sqrt{T}(\hat{\theta}_{ALS} - \hat{\theta}_{GLS}) = o_p(1).$$

# Outline

# The problem

- Consider $X_t = (X'_{1,t}, X'_{2,t})'$ where
  - $X_{1,t}$ is of dimension $d_1 < d$,
  - $X_{2,t}$ of dimension $d_2 = d - d_1$.

# The problem

- Consider $X_t = (X'_{1,t}, X'_{2,t})'$ where
  - $X_{1,t}$ is of dimension $d_1 < d$,
  - $X_{2,t}$ of dimension $d_2 = d - d_1$.

- It is said that $(X_{2,t})$ does not cause (linearly) $(X_{1,t})$ in mean if

$$E(X_{1,t} \mid X_{1,t-1}, \dots) = E(X_{1,t} \mid X_{1,t-1}, X_{2,t-1}, \dots).$$

# The problem

- Consider $X_t = (X'_{1,t}, X'_{2,t})'$ where
  - $X_{1,t}$ is of dimension $d_1 < d$,
  - $X_{2,t}$ of dimension $d_2 = d - d_1$.

- It is said that $(X_{2,t})$ does not cause (linearly) $(X_{1,t})$ in mean if

$$E(X_{1,t} \mid X_{1,t-1}, \dots) = E(X_{1,t} \mid X_{1,t-1}, X_{2,t-1}, \dots).$$

- The problem: check $(X_{2,t})$ does not Granger cause $(X_{1,t})$ in mean.

# The problem

- Consider $X_t = (X'_{1,t}, X'_{2,t})'$ where
  - $X_{1,t}$ is of dimension $d_1 < d$,
  - $X_{2,t}$ of dimension $d_2 = d - d_1$.

- It is said that $(X_{2,t})$ does not cause (linearly) $(X_{1,t})$ in mean if

$$E(X_{1,t} \mid X_{1,t-1}, \dots) = E(X_{1,t} \mid X_{1,t-1}, X_{2,t-1}, \dots).$$

- The problem: check $(X_{2,t})$ does not Granger cause $(X_{1,t})$ in mean.

- In the VAR setup this amounts to test

$$\mathcal{H}_0 : A_{i,12} = 0 \text{ for all } 1 \leq i \leq p,$$

where the $A_{i,12}$'s are the matrices given by the $d_1$ first rows and $d_2$ last columns of the $A_i$'s

# The problem restated

- Let $R$ be a $pd_1 d_2 \times pd^2$ matrix
- Therefore the null hypothesis can be restated

$$\mathcal{H}_0 : R\theta_0 = 0_{pd_1 d_2}.$$

# The problem restated

- Let $R$ be a $pd_1 d_2 \times pd^2$ matrix
- Therefore the null hypothesis can be restated

$$\mathcal{H}_0 : R\theta_0 = 0_{pd_1 d_2}.$$

- Convenient approach: Wald tests.

# Outline

# The standard OLS-based procedure

The commonly used Wald test statistic

$$Q_S = T\hat{\theta}'_{OLS}R'(R\hat{J}^{-1}R')^{-1}R\hat{\theta}_{OLS},$$

with

$$\hat{J} = \left\{T^{-1}\sum_{t=1}^{T}\tilde{X}_{t-1}\tilde{X}'_{t-1}\right\} \otimes \hat{\Omega}_3^{-1}$$

and

$$\hat{\Omega}_3 := T^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}'_t = \Omega_3 + o_p(1),$$

# The modified OLS-based procedure

We propose

$$Q_{OLS} = T\hat{\theta}'_{OLS} R' (R\hat{\Lambda}_3^{-1}\hat{\Lambda}_2\hat{\Lambda}_3^{-1}R')^{-1} R\hat{\theta}_{OLS},$$

where, recall

$$\hat{\Lambda}_2 := T^{-1}\sum_{t=1}^{T}\tilde{X}_{t-1}\tilde{X}'_{t-1}\otimes\hat{u}_t\hat{u}'_t = \Lambda_2 + o_p(1),$$

$$\hat{\Lambda}_3 := T^{-1}\sum_{t=1}^{T}\tilde{X}_{t-1}\tilde{X}'_{t-1}\otimes I_d = \Lambda_3 + o_p(1).$$

## Proposition

*Under suitable assumptions, if $\mathcal{H}_0$ holds true, as $T \to \infty$*

$$Q_{OLS} \Rightarrow \chi^2_{pd_1 d_2},$$

*while*

$$Q_S \Rightarrow Z(\delta) := \sum_{i=1}^{pd_1 d_2} \kappa_i Z_i^2, \tag{2}$$

*where the $Z_i$'s are independent $\mathcal{N}(0,1)$ variables, $\delta = (\kappa_1, \dots, \kappa_{pd_1 d_2})'$ is the vector of the eigenvalues of the matrix*

$$\Psi = (RJ^{-1}R')^{-\frac{1}{2}}(R\Lambda_3^{-1}\Lambda_2\Lambda_3^{-1}R')(RJ^{-1}R')^{-\frac{1}{2}}, \tag{3}$$

*with*

$$J = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p\times p} \otimes \Sigma(r))\tilde{\psi}_i' \right\} dr \otimes \Omega_3^{-1}.$$

# Outline

# The ALS-based Wald procedure

We propose

$$Q_{ALS} = T\hat{\theta}'_{ALS}R'(R\check{\Lambda}_1^{-1}R')^{-1}R\hat{\theta}_{ALS},$$

where, recall

$$\check{\Lambda}_1 = T^{-1}\sum_{t=1}^{T}\tilde{X}_{t-1}\tilde{X}'_{t-1}\otimes\check{\Sigma}_t^{-1} = \Lambda_1 + o_p(1),$$

## Proposition

*Under suitable assumptions, if $\mathcal{H}_0$ holds true, uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ as $T \to \infty$*

$$Q_{ALS} \Rightarrow \chi^2_{pd_1 d_2},$$

## Proposition

*Under suitable assumptions, if $\mathcal{H}_0$ holds true, uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ as $T \to \infty$*

$$Q_{ALS} \Rightarrow \chi^2_{pd_1 d_2},$$

If if $\mathcal{H}_0$ holds true, the (corrected) OLS approach and the ALS approach lead to test statistics with $\chi^2_{pd_1 d_2}$ asymptotic laws.

## Proposition

*Under suitable assumptions, if $\mathcal{H}_0$ holds true, uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ as $T \to \infty$*

$$Q_{ALS} \Rightarrow \chi^2_{pd_1 d_2},$$

If if $\mathcal{H}_0$ holds true, the (corrected) OLS approach and the ALS approach lead to test statistics with $\chi^2_{pd_1 d_2}$ asymptotic laws.

Which one is 'better'?

## Proposition

*Under suitable assumptions, if $\mathcal{H}_0$ holds true, uniformly w.r.t. $b_{kl} \in \mathcal{B}_T$ as $T \to \infty$*

$$Q_{ALS} \Rightarrow \chi^2_{pd_1 d_2},$$

If if $\mathcal{H}_0$ holds true, the (corrected) OLS approach and the ALS approach lead to test statistics with $\chi^2_{pd_1 d_2}$ asymptotic laws.

Which one is 'better'?

## Proposition

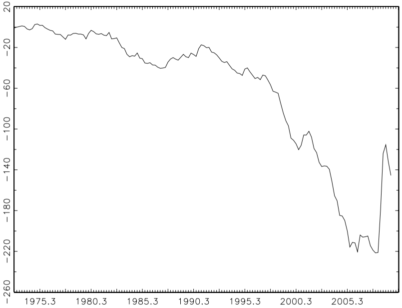
*The relative Bahadur efficiency of the ALS-based test with respect to the OLS-based test is larger or equal to 1 for every fixed alternative.*

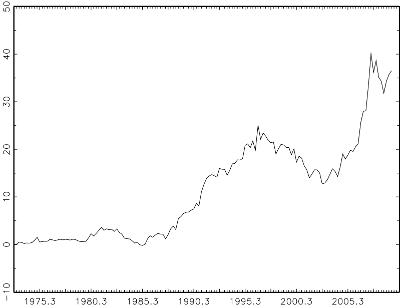
# Outline

1. The model
2. Least squares parameter estimation
3. Adaptive Least Squares parameter estimation
4. Application: testing for linear Granger causality in mean
   - The problem
   - OLS-based Wald tests
   - GLS-based Wald tests
   - Real data illustration
5. Application: portmanteau tests

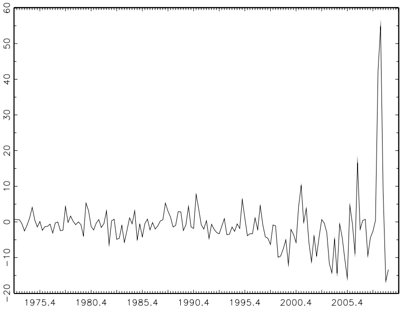
# The data

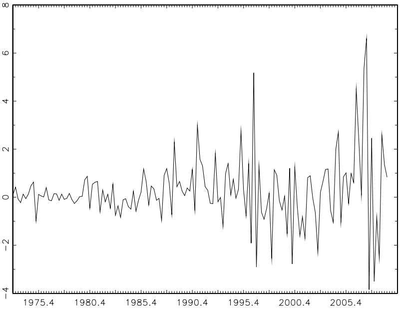
- Quarterly U.S. balance on merchandise trade (left) and balance on services (right) in bln. USD
- Period: from January 1st, 1970 to October 1st, 2009.
- The series are seasonally adjusted



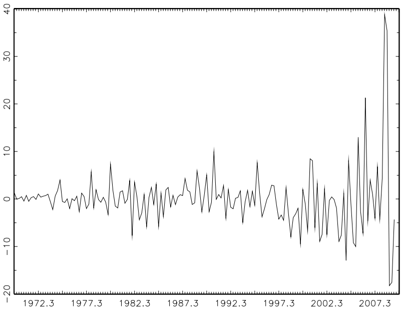



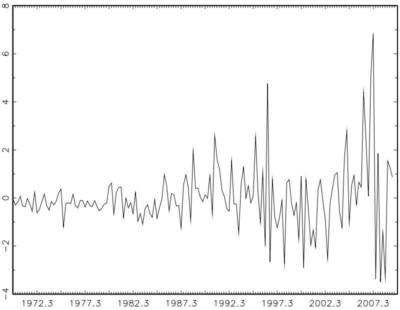
- Presence of unit roots
- Study the series of differences $T = 159$



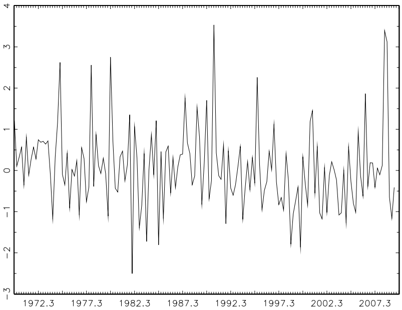



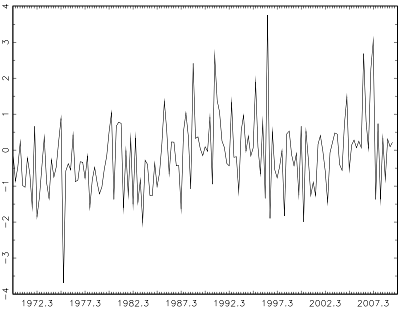
- OLS residuals



- ALS residuals



The p-values of the Wald tests for Granger causality in mean (in %) from the U.S. balance on services to the U.S. balance on merchandize.

| | |
|---|---|
| $W_{OLS}$ | 8.74 |
| $W_S$ | 0.57 |
| $W_{ALS}$ | 25.20 |

# Outline

- Test the order of the multivariate linear VAR(p) model

$$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-p} + u_t$$
$$u_t = H_t \epsilon_t,$$

- Test the order of the multivariate linear VAR(p) model

$$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-p} + u_t$$
$$u_t = H_t \epsilon_t,$$

- Usual way: fix an integer $m > 0$ and test

$$H_0 : \; \mathrm{Cov}(u_t, u_{t-h}) = 0, \text{ for all } 0 < h \leq m,$$

- Test the order of the multivariate linear VAR(p) model

$$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-p} + u_t$$
$$u_t = H_t \epsilon_t,$$

- Usual way: fix an integer $m > 0$ and test

$$H_0 : \ \text{Cov}(u_t, u_{t-h}) = 0, \ \text{for all } \ 0 < h \leq m,$$

The time-varying variance of $u_t$ changes the critical values !!

- Alternative way: test

$$H_0' : \ \mathrm{Cov}(\epsilon_t, \epsilon_{t-h}) = 0, \ \text{for all} \ 0 < h \leq m.$$

- Alternative way: test

$$H_0' : \ \mathrm{Cov}(\epsilon_t, \epsilon_{t-h}) = 0, \ \text{for all} \ 0 < h \leq m.$$

- The values $\epsilon_t$ are approximated using a nonparametric estimate of the deterministic function $H_t$

# Outline

# OLS-based procedure(1/3)

- OLS-based estimates of $u_t$

$$\hat{u}_t = X_t - (\tilde{X}'_{t-1} \otimes I_d)\hat{\theta}_{OLS}$$

# OLS-based procedure(1/3)

- OLS-based estimates of $u_t$

$$\hat{u}_t = X_t - (\tilde{X}'_{t-1} \otimes I_d)\hat{\theta}_{OLS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}^u_{OLS}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{u}_t \hat{u}'_{t-h}.$$

# OLS-based procedure(1/3)

- OLS-based estimates of $u_t$

$$\hat{u}_t = X_t - (\tilde{X}'_{t-1} \otimes I_d)\hat{\theta}_{OLS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}^u_{OLS}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{u}_t \hat{u}'_{t-h}.$$

- The estimates of the first $m$ ($m \geq 1$) residual autocovariances

$$\hat{\gamma}_m^{u,OLS} = \mathrm{vec}\left\{\left(\hat{\Gamma}^u_{OLS}(1), \dots, \hat{\Gamma}^u_{OLS}(m)\right)\right\}$$

# OLS-based procedure (2/3)

- $\Sigma_G = \int_0^1 \Sigma(r)dr$, $\Sigma_{G^{\otimes 2}} = \int_0^1 \Sigma(r)^{\otimes 2}dr$,

$$\Phi_m^u = \sum_{i=0}^{m-1} \left\{ e_m(i+1)e_p(1)' \otimes \Sigma_G \right\} \left\{ K^{i\,\prime} \otimes I_d \right\},$$

$$\Lambda_m^{u,\theta} = \sum_{i=0}^{m-1} \left\{ e_m(i+1)e_p(1)' \otimes \Sigma_{G^{\otimes 2}} \right\} \left\{ (K^{i\,\prime} \otimes I_d \right\},$$

$$\Lambda_m^{u,u} = I_m \otimes \Sigma_{G^{\otimes 2}},$$

where $e_m(j) \in \mathbb{R}^m$ is the vector with $j$th component equal to one and zero elsewhere.

# OLS-based procedure (3/3)

## Proposition

*Under* **A1**,

$$T^{\frac{1}{2}}\hat{\gamma}_m^{u,OLS} \Rightarrow \mathcal{N}(0, \Sigma^{u,OLS}), \tag{4}$$

*where*

$$\Sigma^{u,OLS} = \Lambda_m^{u,u} - \Lambda_m^{u,\theta}\Lambda_3^{-1}\Phi_m^{u\prime} - \Phi_m^u\Lambda_3^{-1}\Lambda_m^{u,\theta\prime} + \Phi_m^u\Lambda_3^{-1}\Lambda_2\Lambda_3^{-1}\Phi_m^{u\prime}, \tag{5}$$

# GLS-based procedure(1/2)

- GLS-based estimates of $\epsilon_t$ using the GLS estimate of the VAR coefficients,

$$\hat{\epsilon}_t = H_t^{-1} X_t - H_t^{-1} (\tilde{X}'_{t-1} \otimes I_d) \hat{\theta}_{GLS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}^{\epsilon}_{GLS}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{\epsilon}_t \hat{\epsilon}'_{t-h}$$

and

$$\hat{\gamma}_m^{\epsilon,GLS} = \text{vec}\left\{ \left( \hat{\Gamma}^{\epsilon}_{GLS}(1), \dots, \hat{\Gamma}^{\epsilon}_{GLS}(m) \right) \right\}.$$

# GLS-based procedure (2/2)

$$\Lambda_m^{\epsilon,\theta} = \sum_{i=0}^{m-1} \left\{ e_m(i+1)e_p(1)' \otimes \int_0^1 \Sigma(r)^{1/2} \otimes \Sigma(r)^{-1/2} dr \right\} \left\{ K^{i\,\prime} \otimes I_d \right\},$$

# GLS-based procedure (2/2)

$$\Lambda_m^{\epsilon,\theta} = \sum_{i=0}^{m-1} \left\{ e_m(i+1)e_p(1)' \otimes \int_0^1 \Sigma(r)^{1/2} \otimes \Sigma(r)^{-1/2} dr \right\} \left\{ K^{i\,\prime} \otimes I_d \right\},$$

## Proposition

*Under* **A1**,

$$T^{\frac{1}{2}} \hat{\gamma}_m^{\epsilon,GLS} \Rightarrow \mathcal{N}(0, \Sigma^{\epsilon,GLS}),$$

*where*

$$\Sigma^{\epsilon,GLS} = I_{d^2 m} - \Lambda_m^{\epsilon,\theta} \Lambda_1^{-1} \Lambda_m^{\epsilon,\theta\,\prime}.$$

# GLS-based procedure (2/2)

$$\Lambda_m^{\epsilon,\theta} = \sum_{i=0}^{m-1} \left\{ e_m(i+1)e_p(1)' \otimes \int_0^1 \Sigma(r)^{1/2} \otimes \Sigma(r)^{-1/2} dr \right\} \left\{ K^{i\,\prime} \otimes I_d \right\},$$

## Proposition

*Under* **A1**,

$$T^{\frac{1}{2}} \hat{\gamma}_m^{\epsilon,GLS} \Rightarrow \mathcal{N}(0, \Sigma^{\epsilon,GLS}),$$

*where*

$$\Sigma^{\epsilon,GLS} = I_{d^2m} - \Lambda_m^{\epsilon,\theta} \Lambda_1^{-1} \Lambda_m^{\epsilon,\theta\,\prime}.$$

Using $\hat{\theta}_{GLS}$ for estimating $\mathrm{Cov}(\epsilon_t, \epsilon_{t-h})$ seems more convenient.

# ALS-based procedure

- ALS-based estimates of $\epsilon_t$,

$$\hat{\epsilon}_t^{ALS} = \hat{H}_t^{-1} X_t - \hat{H}_t^{-1} (\tilde{X}'_{t-1} \otimes I_d) \hat{\theta}_{ALS}$$

# ALS-based procedure

- ALS-based estimates of $\epsilon_t$,

$$\hat{\epsilon}_t^{ALS} = \hat{H}_t^{-1} X_t - \hat{H}_t^{-1} (\tilde{X}_{t-1}' \otimes I_d) \hat{\theta}_{ALS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}_{ALS}^{\epsilon}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{\epsilon}_t^{ALS} \hat{\epsilon}_{t-h}^{ALS\,\prime}$$

# ALS-based procedure

- ALS-based estimates of $\epsilon_t$,

$$\hat{\epsilon}_t^{ALS} = \hat{H}_t^{-1} X_t - \hat{H}_t^{-1} (\tilde{X}'_{t-1} \otimes I_d) \hat{\theta}_{ALS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}^{\epsilon}_{ALS}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{\epsilon}_t^{ALS} \hat{\epsilon}_{t-h}^{ALS\,\prime}$$

## Proposition

*Uniformly w.r.t. $b \in \mathcal{B}_T$*

$$T^{\frac{1}{2}} \left\{ \hat{\Gamma}^{\epsilon}_{ALS}(h) - \hat{\Gamma}^{\epsilon}_{GLS}(h) \right\} = o_p(1), \qquad \forall h.$$

# ALS-based procedure

- ALS-based estimates of $\epsilon_t$,

$$\hat{\epsilon}_t^{ALS} = \hat{H}_t^{-1} X_t - \hat{H}_t^{-1} (\tilde{X}'_{t-1} \otimes I_d) \hat{\theta}_{ALS}$$

- Corresponding residual autocovariances

$$\hat{\Gamma}^{\epsilon}_{ALS}(h) = T^{-1} \sum_{t=h+1}^{T} \hat{\epsilon}_t^{ALS} \hat{\epsilon}_{t-h}^{ALS\,\prime}$$

## Proposition

*Uniformly w.r.t.* $b \in \mathcal{B}_T$

$$T^{\frac{1}{2}} \left\{ \hat{\Gamma}^{\epsilon}_{ALS}(h) - \hat{\Gamma}^{\epsilon}_{GLS}(h) \right\} = o_p(1), \qquad \forall h.$$

GLS and ALS versions of autocovariances and autocorrelations estimates are asymptotically equivalent

# Outline

- Box-Pierce statistic based on OLS estimates of $\mathrm{Cov}(u_t, u_{t-h})$

$$Q_m^{OLS} = T \sum_{h=1}^{m} \mathrm{tr}\left(\hat{\Gamma}^{u\prime}_{OLS}(h)\hat{\Gamma}^{u}_{OLS}(0)^{-1}\hat{\Gamma}^{u}_{OLS}(h)\hat{\Gamma}^{u}_{OLS}(0)^{-1}\right)$$

- Box-Pierce statistic based on ALS estimates of $\mathrm{Cov}(\epsilon_t, \epsilon_{t-h})$

$$Q_m^{ALS} = T \sum_{h=1}^{m} \mathrm{tr}\left(\hat{\Gamma}^{\epsilon\prime}_{ALS}(h)\hat{\Gamma}^{\epsilon}_{ALS}(0)^{-1}\hat{\Gamma}^{\epsilon}_{ALS}(h)\hat{\Gamma}^{\epsilon}_{ALS}(0)^{-1}\right)$$

# Asymptotics for OLS-based Box-Pierce statistic

## Proposition

*The statistic $Q_m^{OLS}$ converges in law to*

$$U(\delta_m^{OLS}) = \sum_{i=1}^{d^2 m} \delta_i^{ols} U_i^2, \tag{6}$$

*as $T \to \infty$, where $\delta_m^{OLS} = (\delta_1^{ols}, \dots, \delta_{d^2 m}^{ols})'$ is the vector of the eigenvalues of the matrix*

$$\Delta_m^{OLS} = (I_m \otimes \Sigma_G^{-1/2} \otimes \Sigma_G^{-1/2}) \Sigma^{u,OLS} (I_m \otimes \Sigma_G^{-1/2} \otimes \Sigma_G^{-1/2}),$$

$\Sigma_G = \int_0^1 \Sigma(r) dr$ *and the $U_i$'s are independent $\mathcal{N}(0,1)$ variables.*

# Asymptotics for ALS-based Box-Pierce statistic

## Proposition

*The statistic $Q_m^{ALS}$ converges in law to*

$$U(\delta_m^{ALS}) = \sum_{i=1}^{d^2 m} \delta_i^{als} U_i^2,$$

*as $T \to \infty$, where $\delta_m^{ALS} = (\delta_1^{als}, \dots, \delta_{d^2 m}^{als})'$ is the vector of the eigenvalues of $\Sigma^{\epsilon, GLS}$, and the $U_i$'s are independent $\mathcal{N}(0,1)$ variables.*

# Asymptotics for ALS-based Box-Pierce statistic

**Proposition**

*The statistic $Q_m^{ALS}$ converges in law to*

$$U(\delta_m^{ALS}) = \sum_{i=1}^{d^2 m} \delta_i^{als} U_i^2,$$

*as $T \to \infty$, where $\delta_m^{ALS} = (\delta_1^{als}, \dots, \delta_{d^2 m}^{als})'$ is the vector of the eigenvalues of $\Sigma^{\epsilon, GLS}$, and the $U_i$'s are independent $\mathcal{N}(0,1)$ variables.*

The classical Khi-square limit law is recovered with the ALS approach when $\Sigma(r) = \sigma^2(r) I_d$.

Merci pour votre attention!