Inference of high-dimensional VAR models

Seminar:

"Modeling, Simulation and Inference of Complex Biological Systems"

Katharina Schneider

07.07.2006

Table of contents

Stochastic process

Sequence of random variables $Y = \{Y_t, t \in T\}$

- Trend: $\mu_t = \mathbb{E}(Y_t)$
- Variance: $\sigma_t = \mathbb{E}[(Y_t - \mu_t)^2]$
- Autocovariance: $\gamma_{t,s} = \mathbb{E}\{[Y_t - \mu_t][Y_s - \mu_s]\}$
- Stationarity
  - Y_t strongly stationary :⇔
    - $\forall n, t_1, \ldots, t_n, h$:
      $F_{x_{t_1},\ldots,x_{t_n}}(x_1,\ldots,x_n) = F_{x_{t_1+h},\ldots,x_{t_n+h}}(x_1,\ldots,x_n)$
  - Y_t weakly stationary :⇔
    - $\mu_t = \mu = const$
    - $\sigma_t^2 = \sigma^2 = const$
    - $\gamma_{t,s} = \gamma_{t-s} = \gamma_k$ with $k = t - s$ (Lag)
- Autocovariance function: $\gamma_k = \mathbb{E}\{[Y_t - \mu][Y_{t-k} - \mu]\}$
- Autocorrelation function (ACF): $\varrho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma^2}$
  (by standardization with $\sigma^2 = \gamma_0$)

Linear time series and stationarity

One finite realization of a stochastic process $y = \{y_t, t \in T\}$
Classical decomposition model: $y_t = \mu_t + s_t + u_t$
(= trend + seasonal component + stationary random noise)

- Stationarity
  Y_t stationary ⇔ y_t stationary
- descriptive analysis of stationarity with graphs and correlograms:
  - no trend
  - no systematic change of variance
  - no strictly periodic fluctuations
- Tests on stationarity:
  - Unit Root Tests (Dickey-Fuller-Test, Augmented DF-Test)
- approaches to obtain stationarity: differentiation, integration, filtering

Basics

- Lag operator L (Backshift operator)

$$L^0 y_t = y_t \qquad L^1 y_t = y_{t-1} \qquad L^2 y_t = y_{t-2} \qquad \ldots \qquad \ldots \qquad L^k y_t = y_{t-k}$$

- White noise ϵ_t
  - a series of iid random variables ("innovations", "shocks")
  - $\mathbb{E}(\epsilon_t) = \mu_t = 0$
  - σ_ϵ^2 $(\mathbf{\Sigma}_\epsilon)$
  - $\gamma_{t,s} = 0$ for $t \neq s$
- Properties ACF
  - $\varrho(k) = \varrho(-k)$
  - $-1 \leq \varrho(k) \leq 1$
  - $Y(t)$ and $Y(t-k)$ independent ⇒ $\varrho(k) = 0$

  Correlogram: graph of ϱ

Linear time series models

	univariate	multivariate
stationary	AR MA ARMA	VAR VMA VARMA
non-stationary	ARIMA	VARIMA

Modeling a time series

1. diagnosis (stationarity, autocorrelation, etc.)
2. model identification
   - d: order of integration $\widehat{=}$ number of differentiations for stationarity
   - p, q: with Box-Jenkins(ACF, PACF, etc.), AIC, Bayes-Schwarz, etc.
3. estimation of the parameters (LS, ML, etc.)
4. model selection

Autoregressive process of order p

- **AR(p)**

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t \quad \Leftrightarrow \quad \mathbf{\Phi}(L) y_t = \epsilon_t$$

- Properties
  - $\mathbb{E}(y_t) = 0$
  - $\mathbb{V}ar(y_t) = const.$
  - $\left.\begin{array}{ll} \gamma_k = \sum_{l=1}^{p} \phi_l \gamma_{k-l} & : k = 1, 2, \ldots \\ \varrho_k = \sum_{l=1}^{p} \phi_l \varrho_{k-l} & : k = 1, 2, \ldots \end{array}\right\} Yule-Walker$
- Stationarity
  Characteristic equation: $\mathbf{\Phi}(u) = 0$ with $u \in \mathbb{C}$
  - AR(p) stationary: $|u| > 1$ ↔ if all (complex) solutions of the characteristic equation lie outside the unit circle
  - AR(p) nonstationary: $|u| = 1$ (unit root)

Moving average process of order q

- **MA(q)**

$$y_t = \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \quad \Leftrightarrow \quad y_t = \mathbf{\Theta}(L)\epsilon_t$$

- Properties
  - $\mathbb{E}(y_t) = 0$
  - $\mathbb{V}ar(y_t) = \sigma^2 \sum_{i=0}^{q} \theta_i^2$
  - $\gamma_k = \begin{cases} \sigma^2 \sum_{i=0}^{q-k} \theta_{i+k}\theta_i & : k = 0, 1, \ldots, q \\ 0 & : k > q \end{cases}$
- Stationarity
  $\mathbb{E}(y_t), \mathbb{V}ar(y_t), \gamma_k$ independent of t ⇒ MA(q) weakly stationary

Autoregressive moving average process

- **ARMA(p,q)**

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \quad \Leftrightarrow \quad \mathbf{\Phi}(L) y_t = \mathbf{\Theta}(L)\epsilon_t$$

  - Stationarity
    ARMA(p,q) stationary $\Leftrightarrow$ AR(p)-part stationary
- **ARIMA(p,d,q)**
  Autoregressive Integrated Moving Average Process

$$\mathbf{\Phi}(L)(1-L)^d y_t = \mathbf{\Theta}(L)\epsilon_t$$

  $\Leftrightarrow$ if $x_t := (1-L)^d y_t$ ARMA(p,q)$\rightarrow \{y_t\}$ ARIMA(p,d,q)
  - Stationarity
    ARIMA(p,d,q) stationary $\Leftrightarrow$ $d = 0$ (i.e. ARMA(p,q))

Vector autoregressive model of order p

VAR(p)

$$\mathbf{y}'_t = \mathbf{c} + \sum_{i=1}^{L} \mathbf{y}'_{t-i}\mathbf{B}_i + \boldsymbol{\epsilon}'_t \qquad \text{for } t = 1, \ldots, T$$

with

$\mathbf{y}_t$	$(1 \times p)$ random vector
$\mathbf{c}$	unknown fixed $(1 \times p)$ vector of intercept terms
$\mathbf{B}_i$	unknown fixed $(p \times p)$ regression coefficient matrices
$\boldsymbol{\epsilon}_t$	p-dimensional white noise process ($\boldsymbol{\epsilon}_t \sim iid(\mathbf{0}, \mathbf{\Sigma})$)
L	known positive integer (number of lags)
t	time period variable

Vector autoregressive model of order p

VAR(p)

$$\mathbf{Y} = \mathbf{X}\mathbf{\Phi} + \boldsymbol{\epsilon} = \underbrace{\begin{pmatrix} \mathbf{x}'_1 \\ \vdots \\ \vdots \\ \mathbf{x}'_T \end{pmatrix}}_{T\times(1+Lp)} \underbrace{\begin{pmatrix} \mathbf{c} \\ \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_L \end{pmatrix}}_{(1+Lp)\times p} + \underbrace{\begin{pmatrix} \boldsymbol{\epsilon}'_1 \\ \vdots \\ \vdots \\ \boldsymbol{\epsilon}'_T \end{pmatrix}}_{T\times p} = \underbrace{\begin{pmatrix} \mathbf{y}'_1 \\ \vdots \\ \vdots \\ \mathbf{y}'_T \end{pmatrix}}_{T\times p}$$

with

$$\mathbf{x}_t = \underbrace{\begin{pmatrix} 1 \\ \mathbf{y}'_{t-1} \\ \vdots \\ \mathbf{y}'_{t-L} \end{pmatrix}}_{(1+Lp)\times 1} \qquad \begin{aligned} \epsilon_t \sim \ & iid(\mathbf{0}, \mathbf{\Sigma}) \\ & \mathbf{\Sigma} := \mathbb{E}(\epsilon_t \epsilon'_t) = Cov(\epsilon_t) \\ & \ : \ \text{positive definite } p \times p \text{ matrix} \end{aligned}$$

Vector autoregressive model of order p

Stability
$VAR(p)$ stable $\Leftrightarrow$ $det(\mathbf{I} - \mathbf{B}u) \neq 0$ for $|u| \leq 1$

Stationarity
$VAR(p)$ stable $\Rightarrow$ $VAR(p)$ stationary

- $L = 1$ (VARs with one lag):
  - $VAR(p)$ stationary $\Leftrightarrow$ the absolute values of the real eigenvalues of $\mathbf{B}_1$ are less than unity
  - $VAR(p)$ nonstationary $\Leftrightarrow$ the absolute values of the real eigenvalues of $\mathbf{B}_1$ lie on the unit circle
- $L > 1$ (VARs with more than one lag):
  $\Rightarrow$ rewrite as a VAR with one lag

$\Rightarrow$ stationarity is necessary for impulse response analysis and for estimation
$\Rightarrow$ problem: differentiation sometimes $\rightsquigarrow$ falsification

Moving average representation of a VAR(p)

$$\mathbf{y}'_t = E_0\mathbf{y}'_t + \sum_{j=0}^{t-1} \boldsymbol{\epsilon}'_{t-j}\mathbf{H}_j$$

with

VAR	stationary
$\mathbf{H}_0$	$(p \times p)$ identity matrix
$\mathbf{H}_j$	impulse responses to a shock occuring j periods ago

$\Rightarrow$ $\mathbf{y}_t$ is expressed in terms of past and present error/innovation vectors $\boldsymbol{\epsilon}_t$ and the mean term

$\Rightarrow$ necessary for impulse response analysis

$\Rightarrow$ can be used to determine the autocovariances

Characterization and interpretation of VARs

Characteristics

- VAR models are popular tools for analyzing multivariate time series data
- VAR models represent the correlations among a set of variables
  $\Rightarrow$ analysis of certain aspects of the relationships between the interesting variables

Structural analysis

- Granger-Causality
- impulse response analysis

Granger-Causality

- based on the principle of cause and effect
- MA representation of a $K-$ dimensional VAR

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{H}(L)\boldsymbol{\epsilon}_t \quad \text{with} \quad \mathbf{H}_0 = \mathbf{I}_K$$

- partitioned MA representation

$$\mathbf{y}_t = \begin{pmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \mathbf{H}_{11}(L) & \mathbf{H}_{12}(L) \\ \mathbf{H}_{21}(L) & \mathbf{H}_{22}(L) \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

  with
  $\mathbf{z}_t$ M-dimensional
  $\mathbf{x}_t$ $(K-M)$-dimensional

  $\Rightarrow \mathbf{z}_t$ is not Granger-caused by $\mathbf{x}_t$ $:\Leftrightarrow$ $\mathbf{H}_{12,i} = 0$ for $i = 1, 2, \ldots$
- $\mathbf{x}_t$ is Granger-causal to $\mathbf{z}_t$ $\Leftrightarrow$ the information in the past and the present of $\mathbf{x}$ helps to predict $\mathbf{z}_{t+1}$

Impulse response functions

- MA representation

$$\mathbf{y}'_t = E_0\mathbf{y}'_t + \sum_{j=0}^{t-1} \boldsymbol{\epsilon}'_{t-j}\mathbf{H}_j$$

- impulse responses of $\mathbf{y}_t$ to a shock $\boldsymbol{\epsilon}_{t-j}$ occuring j periods earlier

$$\mathbf{H}_j = \sum_{i=1}^{j} \mathbf{B}_i \mathbf{H}_{j-i}$$

  with
  $\mathbf{B}_i = 0$ for $i > L$
  $\boldsymbol{\epsilon}_t$ correlated

Problem: ϵ_t are correlated $\longrightarrow$ identification problem
Solution: orthogonalization of the errors

- Cholesky decomposition of the covariance matrix

$$\boldsymbol{\Sigma} = \boldsymbol{\Psi}'\boldsymbol{\Psi}$$

with

$\boldsymbol{\Psi}$ uppertriangular positive definite matrix

- connection between structural shocks and VAR errors

$$\mathbf{u}_t' = \epsilon_t'\boldsymbol{\Psi}^{-1}$$

with

$\mathbf{u}_t$ structural error vector (with $\boldsymbol{\Sigma}(\mathbf{u}_t)$: identity matrix)

- impulse responses to structural shocks occuring j periods earlier

$$\mathbf{Z}_j = \boldsymbol{\Psi}\mathbf{H}_j$$

Possibilities for estimating a VAR

- Least Squares
  $\Rightarrow$ asymptotic properties
- Maximum Likelihood
  $\Rightarrow$ assumption: known distribution of data
  $\Rightarrow$ for some distributions MLE does not have an analytical form or does not exist
- Bayes
  $\Rightarrow$ effectiveness for finite-sample inference
  $\Rightarrow$ the estimated process may be used for prediction and economic analyses

$\Rightarrow$ applicable to estimate the model parameters and to estimate the distributions of the impulse response functions

MLE for normal VARs

- $\mathbf{Y} = \mathbf{X}\boldsymbol{\Phi} + \epsilon \qquad \epsilon \overset{iid}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma})$
- Likelihood function of $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$

$$\begin{aligned} l_N(\boldsymbol{\Phi}, \boldsymbol{\Sigma}) &= \frac{1}{|\boldsymbol{\Sigma}|^{T/2}}\, exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\Phi})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \mathbf{x}_t\boldsymbol{\Phi})\right\} \\ &= \frac{1}{|\boldsymbol{\Sigma}|^{T/2}}\, \underbrace{etr}_{=exp(trace)}\left\{-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi})\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi})'\right\} \end{aligned}$$

- Maximum likelihood estimators (MLEs)

$$\begin{aligned} \widehat{\boldsymbol{\Phi}}_{MLE} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\ \widehat{\boldsymbol{\Sigma}}_{MLE} &= \frac{\mathbf{S}(\widehat{\boldsymbol{\Phi}}_{MLE})}{T} \qquad \text{with} \qquad \mathbf{S}(\boldsymbol{\Phi}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi}) \end{aligned}$$

MLE for Student-t VARs

- $\mathbf{Y} = \mathbf{X}\boldsymbol{\Phi} + \epsilon \qquad \epsilon_t \overset{ind.}{\sim} t_\nu(\mathbf{0}, \boldsymbol{\Sigma})$
- Density of $t_\nu(\mathbf{0}, \boldsymbol{\Sigma})$ (multivariate-t distribution)

$$p\left(\mathbf{s}|\boldsymbol{\Sigma}, \nu\right) = \frac{\Gamma\left(\frac{1}{2}(\nu + p)\right)}{(\pi\nu)^{p/2}\,\Gamma\left(\frac{\nu}{2}\right)} \times |\boldsymbol{\Sigma}|^{-1/2}\left(1 + \frac{1}{\nu}\mathbf{s}'\,\boldsymbol{\Sigma}^{-1}\mathbf{s}\right)^{\frac{-(p+\nu)}{2}}, \mathbf{s} \in \mathbb{R}^p$$

- Maximum likelihood estimators (MLEs)
  - ν given $\Rightarrow$ MLE for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$ is not available in closed form
  - ν unknown $\Rightarrow$ MLE for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma}, \nu)$ may not even exist

Basics

A Bayesian estimator of $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$ depends on

- the distribution model
- the prior
- the loss function

Bayesian procedure

1. choose a prior
2. derive/compute the posterior
3. choose a loss function
4. estimate under the loss function
5. calculate the risk function
6. evaluate the performance of the estimates

Priors for $\phi, \boldsymbol{\Sigma}$

for $\boldsymbol{\Sigma}$ / for ϕ	Jeffreys prior	RATS prior	Reference prior
Constant prior	$\pi_{CJ}(\phi, \boldsymbol{\Sigma})$	$\pi_{CA}(\phi, \boldsymbol{\Sigma})$	$\pi_{CR}(\phi, \boldsymbol{\Sigma})$
Shrinkage prior	$\pi_{SJ}(\phi, \boldsymbol{\Sigma})$	$\pi_{SA}(\phi, \boldsymbol{\Sigma})$	$\pi_{SR}(\phi, \boldsymbol{\Sigma})$

$\Rightarrow$ noninformative priors

Prior for ν in the Student-t VAR

$$w = \frac{\nu}{2} \quad \text{with} \quad w \sim Gamma(a, b)$$

a, b known positive constants

Jeffreys prior

Jeffreys prior for the normal VAR model

- Jeffreys prior for $\boldsymbol{\Sigma}$: $\pi_J(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(p+1)/2}$
- constant Jeffreys prior for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$: $\pi_{CJ}(\phi, \boldsymbol{\Sigma}) \propto \pi_J(\boldsymbol{\Sigma})$
  - conditional posterior of ϕ given $(\boldsymbol{\Sigma}, \mathbf{Y})$:

$$N_J(\widehat{\phi}_{MLE}, \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1})$$

  - marginal posterior of $\boldsymbol{\Sigma}$ given $\mathbf{Y}$:

$$Inverse\ Wishart(\mathbf{S}(\widehat{\boldsymbol{\Phi}}_{MLE}),\ T - Lp - 1)$$

$\Rightarrow$ derived from the "invariance principle"

- Shrinkage Jeffreys prior for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$: $\pi_{SJ}(\phi, \boldsymbol{\Sigma}) = \pi_S(\phi)\pi_J(\boldsymbol{\Sigma})$
  $\Rightarrow$ motivated by Stein´s result on inadmissibility of the MLE

Loss functions for $\boldsymbol{\Sigma}$

1. pseudoentropy loss

$$L_{\boldsymbol{\Sigma}1}(\widehat{\boldsymbol{\Sigma}}; \boldsymbol{\Sigma}) = trace(\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}) - log|\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}| - p$$

2. quadratic loss

$$L_{\boldsymbol{\Sigma}2}(\widehat{\boldsymbol{\Sigma}}; \boldsymbol{\Sigma}) = trace(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \mathbf{I})^2$$

3. pseudoentropy function on $\boldsymbol{\Sigma}^{-1}$

$$L_{\boldsymbol{\Sigma}3}(\widehat{\boldsymbol{\Sigma}}; \boldsymbol{\Sigma}) = trace(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) - log|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}| - p$$

Loss functions for $\mathbf{\Phi}$

1. quadratic loss

$$L_{\mathbf{\Phi}1}(\widehat{\mathbf{\Phi}}, \mathbf{\Phi}) = trace\left\{(\widehat{\mathbf{\Phi}} - \mathbf{\Phi})' \, \mathbf{W} \, (\widehat{\mathbf{\Phi}} - \mathbf{\Phi})\right\}$$

with $\mathbf{W}$ constant weighting matrix

$$\text{here: } \mathbf{W} = \mathbf{I} \quad \Rightarrow \quad L_{\mathbf{\Phi}1} = \sum_{i=1}^{1+Lp} \sum_{j=1}^{p} (\widehat{\phi}_{ij} - \phi_{ij})^2$$

$\rightarrow$ symmetric

2. LINEX loss

$$L_{\mathbf{\Phi}2}(\widehat{\mathbf{\Phi}}, \mathbf{\Phi}) = \sum_{i=1}^{1+Lp} \sum_{j=1}^{p} \left\{ exp\left[a_{ij}(\widehat{\phi}_{ij} - \phi_{ij})\right] - a_{ij}(\widehat{\phi}_{ij} - \phi_{ij}) - 1 \right\}$$

with a_{ij} given constant
$\rightarrow$ asymmetric

Loss and risk function for impulse response functions

Loss function

$$L(\mathbf{Z}_j, \widehat{\mathbf{Z}}_j) = trace\left\{(\mathbf{Z}_j - \widehat{\mathbf{Z}}_j)' \, \mathbf{\Omega} \, (\mathbf{Z}_j - \widehat{\mathbf{Z}}_j)\right\}$$

with $\mathbf{\Omega}$: weighting matrix for the estimation error of each element of the impulse responses
$\rightarrow$ may be determined by the economic significance of the element (here: identity matrix)

Risk function

$$R_{Imp,i} = \frac{1}{N} \sum_{n=1}^{N} trace\left\{\left(\mathbf{Z}_i - \widehat{\mathbf{Z}}_i^{(n)}\right)' - \left(\mathbf{Z}_i - \widehat{\mathbf{Z}}_i^{(n)}\right)\right\}$$

with $\widehat{\mathbf{Z}}_i^{(n)}$: impulse response matrix for the ith step after the shock for the nth dataset generated in the experiment

Simulations and Bayesian computation

1. generate $N = 1,000$ data samples from a VAR(5) model with one lag ($L = 1$) and the known parameters

$$\mathbf{\Sigma} = \begin{pmatrix} 1.0 & & \mathbf{0.5} \\ & \ddots & \\ \mathbf{0.5} & & 1.0 \end{pmatrix}, \quad \mathbf{\Phi} = \begin{pmatrix} \mathbf{c} \\ \mathbf{B}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{I}_5 \end{pmatrix}, \quad \begin{array}{l} \nu = 8 \\ \text{prior for } \nu: \nu \sim Gamma(1, 0.5) \end{array}$$

2. compute the Bayesian estimates under competing priors and the different losses via MCMC ($M = 10,000$ cycles)
   1. find the full conditional distributions of $(\phi, \mathbf{\Sigma})$ with $\phi = vec(\mathbf{\Phi})$
   2. simulate the posteriors of $(\mathbf{\Phi}, \mathbf{\Sigma})$
3. estimate the frequentist risks under a loss L of the estimates as the average loss belonging to $\widehat{\mathbf{\Sigma}}$ and $\widehat{\mathbf{\Phi}}$ across generated data samples
4. evaluate the performance of the Bayesian estimates in terms of the frequentist risks given the true parameters

Frequentist average losses of competing Bayes estimates of $\mathbf{\Sigma}$ in the normal VAR

	$L_{\Sigma 1}$	$L_{\Sigma 2}$	$L_{\Sigma 3}$
$\widehat{\Sigma}_{MLE}$	.861 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{1CA}$	.608 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{2CA}$	1.187 (.501)	.803 (.192)	.660 (.205)
$\widehat{\Sigma}_{3CA}$	.861 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{1CJ}$	.450 (.222)	.800 (.344)	.389 (.142)
$\widehat{\Sigma}_{2CJ}$	.862 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{3CJ}$	.609 (.309)	.645 (.222)	.415 (.153)
$\widehat{\Sigma}_{1CR}$	.281 (.172)	.434 (.219)	.234 (.113)
$\widehat{\Sigma}_{2CR}$	.546 (.314)	.489 (.184)	.353 (.161)
$\widehat{\Sigma}_{3CR}$	.386 (.238)	.419 (.178)	.273 (.133)
$\widehat{\Sigma}_{1SA}$	.609 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{2SA}$	1.187 (.501)	.803 (.192)	.661 (.205)
$\widehat{\Sigma}_{3SA}$	.862 (.403)	.682 (.189)	.516 (.178)
$\widehat{\Sigma}_{1SJ}$	.449 (.221)	.801 (.345)	.389 (.143)
$\widehat{\Sigma}_{2SJ}$	.862 (.403)	.682 (.189)	.516 (.178)
$\widehat{\Sigma}_{3SJ}$	.609 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{1SR}$	.261 (.163)	.418 (.208)	.221 (.107)
$\widehat{\Sigma}_{2SR}$	.505 (.300)	.464 (.177)	.332 (.154)
$\widehat{\Sigma}_{3SR}$	.356 (.226)	.399 (.169)	.256 (.126)

Frequentist average losses of competing Bayes estimates of $\mathbf{\Sigma}$ in the normal VAR

	$L_{\Sigma 1}$	$L_{\Sigma 2}$	$L_{\Sigma 3}$
$\widehat{\Sigma}_{MLE}$	.861 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{1CA}$	.608 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{2CA}$	1.187 (.501)	.803 (.192)	.660 (.205)
$\widehat{\Sigma}_{3CA}$	.861 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{1CJ}$	.450 (.222)	.800 (.344)	.389 (.142)
$\widehat{\Sigma}_{2CJ}$	.862 (.403)	.681 (.189)	.516 (.178)
$\widehat{\Sigma}_{3CJ}$	.609 (.309)	.645 (.222)	.415 (.153)
$\widehat{\Sigma}_{1CR}$	.281 (.172)	.434 (.219)	.234 (.113)
$\widehat{\Sigma}_{2CR}$	.546 (.314)	.489 (.184)	.353 (.161)
$\widehat{\Sigma}_{3CR}$	.386 (.238)	.419 (.178)	.273 (.133)
$\widehat{\Sigma}_{1SA}$	.609 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{2SA}$	1.187 (.501)	.803 (.192)	.661 (.205)
$\widehat{\Sigma}_{3SA}$	.862 (.403)	.682 (.189)	.516 (.178)
$\widehat{\Sigma}_{1SJ}$	.449 (.221)	.801 (.345)	.389 (.143)
$\widehat{\Sigma}_{2SJ}$	.862 (.403)	.682 (.189)	.516 (.178)
$\widehat{\Sigma}_{3SJ}$	.609 (.308)	.646 (.222)	.415 (.153)
$\widehat{\Sigma}_{1SR}$	.261 (.163)	.418 (.208)	.221 (.107)
$\widehat{\Sigma}_{2SR}$	.505 (.300)	.464 (.177)	.332 (.154)
$\widehat{\Sigma}_{3SR}$	.356 (.226)	.399 (.169)	.256 (.126)

Frequentist average losses of competing Bayes estimates of $\mathbf{\Phi}$ in the normal VAR

	$L_{\Phi 1}$	$L_{\Phi 2}$
$\widehat{\Phi}_{MLE}$	11.183 (14.070)	11.614 (6.898)
$\widehat{\Phi}_{1CA}$	11.184 (14.068)	11.611 (6.893)
$\widehat{\Phi}_{2CA}$	11.135 (14.055)	9.416 (5.154)
$\widehat{\Phi}_{1CJ}$	11.184 (14.094)	11.613 (6.905)
$\widehat{\Phi}_{2CJ}$	11.134 (14.079)	9.156 (4.947)
$\widehat{\Phi}_{1CR}$	11.185 (14.070)	11.615 (6.894)
$\widehat{\Phi}_{2CR}$	11.135 (14.057)	9.319 (5.066)
$\widehat{\Phi}_{1SA}$	1.552 (.597)	7.254 (3.349)
$\widehat{\Phi}_{2SA}$	1.523 (.596)	6.313 (2.786)
$\widehat{\Phi}_{1SJ}$	1.419 (.501)	7.174 (3.250)
$\widehat{\Phi}_{2SJ}$	1.387 (.499)	6.155 (2.653)
$\widehat{\Phi}_{1SR}$	1.261 (.359)	7.176 (3.332)
$\widehat{\Phi}_{2SR}$	1.231 (.355)	6.198 (2.739)

Frequentist average losses of competing Bayes estimates of $\mathbf{\Phi}$ in the normal VAR

	$L_{\Phi 1}$	$L_{\Phi 2}$
$\widehat{\Phi}_{MLE}$	11.183 (14.070)	11.614 (6.898)
$\widehat{\Phi}_{1CA}$	11.184 (14.068)	11.611 (6.893)
$\widehat{\Phi}_{2CA}$	11.135 (14.055)	9.416 (5.154)
$\widehat{\Phi}_{1CJ}$	11.184 (14.094)	11.613 (6.905)
$\widehat{\Phi}_{2CJ}$	11.134 (14.079)	9.156 (4.947)
$\widehat{\Phi}_{1CR}$	11.185 (14.070)	11.615 (6.894)
$\widehat{\Phi}_{2CR}$	11.135 (14.057)	9.319 (5.066)
$\widehat{\Phi}_{1SA}$	1.552 (.597)	7.254 (3.349)
$\widehat{\Phi}_{2SA}$	1.523 (.596)	6.313 (2.786)
$\widehat{\Phi}_{1SJ}$	1.419 (.501)	7.174 (3.250)
$\widehat{\Phi}_{2SJ}$	1.387 (.499)	6.155 (2.653)
$\widehat{\Phi}_{1SR}$	1.261 (.359)	7.176 (3.332)
$\widehat{\Phi}_{2SR}$	1.231 (.355)	6.198 (2.739)

Estimation of impulse response functions

$\mathbf{Z}_j$ are nonlinear functions of $(\mathbf{\Phi}, \mathbf{\Sigma})\Rightarrow$ Bayes-simulations

1. generate data samples from a VAR with known parameters
2. compute the Bayesian estimates under competing priors via MCMC
   1. find the full conditional distributions of $(\mathbf{Z})$
   2. simulate the posteriors of $(\mathbf{Z})$
   3. compute $\widehat{\mathbf{Z}}_j = \mathbb{E}(\mathbf{Z}_j|\,\mathbf{Y})$ (assumption: $\mathbf{\Omega}$ const.)
3. evaluate the performance of the estimates in terms of the frequentist average of sum of squared errors

Frequentist average losses of impulse responses in the normal VAR

Horizon	MLE	CA	CJ	CR	SA	SJ	SR
1	.963	.911	.848	.854	.734	.662	.640
2	1.745	1.657	1.603	1.640	1.331	1.249	1.250
3	2.389	2.276	2.238	2.290	1.851	1.772	1.781
4	2.898	2.767	2.747	2.804	2.282	2.208	2.220
5	3.302	3.158	3.152	3.210	2.638	2.568	2.580
6	3.625	3.470	3.476	3.533	2.931	2.867	2.877
7	3.886	3.723	3.738	3.793	3.176	3.116	3.125
8	4.101	3.929	3.953	4.005	3.382	3.325	3.332
9	4.278	4.101	4.131	4.180	3.556	3.502	3.508
10	4.427	4.246	4.283	4.330	3.706	3.653	3.659
11	4.554	4.371	4.417	4.460	3.835	3.785	3.789
12	4.662	4.483	4.540	4.579	3.947	3.900	3.903

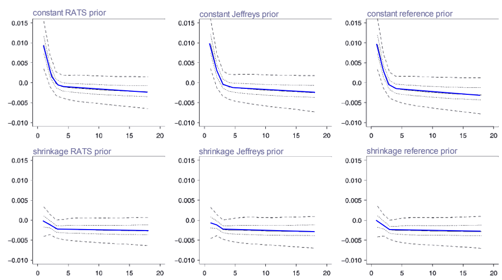
Frequentist average losses of impulse responses in the Student-t VAR

Horizon	CA	CJ	CR	SA	SJ	SR
1	.930	.867	.875	.747	.673	.654
2	1.684	1.630	1.670	1.334	1.252	1.255
3	2.307	2.272	2.327	1.843	1.764	1.776
4	2.798	2.782	2.842	2.264	2.190	2.206
5	3.186	3.184	3.246	2.610	2.541	2.558
6	3.492	3.502	3.564	2.895	2.831	2.849
7	3.737	3.756	3.816	3.133	3.073	3.090
8	3.934	3.960	4.019	3.333	3.276	3.293
9	4.095	4.127	4.184	3.502	3.448	3.465
10	4.229	4.266	4.321	3.646	3.595	3.612
11	4.342	4.386	4.438	3.771	3.721	3.738
12	4.441	4.493	4.542	3.879	3.832	3.848

VAR of U.S. Economy

- VAR(6) model with two lags (L=2)
- quarterly data of the U.S. economy from 1959 Q1 to 2001 Q4
  - real GDP(Gross Domestic Product)
  - GDP deflator
  - world commodity price
  - Federal Funds rates
  - nonborrowed reserves
  - M2 money stock
- $M = 10,000$ MCMC cycles

Responses of GDP to an inflation shock



Concluding remarks

Simulations

→ the choice of prior has stronger effects on the Bayesian estimates than the choice of loss function

→ the asymmetric LINEX estimator for $\mathbf{\Phi}$ does better overall than the posterior mean

→ there is no estimator for $\mathbf{\Sigma}$ dominating in all cases

→ the shrinkage prior dominates the constant prior

→ reference prior on $\mathbf{\Sigma}$ dominates the Jeffreys prior and the RATS prior

Concluding remarks

VAR of U.S. Economy

→ significant improvement of the estimates by using alternative priors in place of constant prior

→ impulse responses of GDP to an inflation shock are distinctly different under the competing priors

→ the posterior losses under the shrinkage reference prior are smaller

→ VAR model estimates allow some degree of collinearity and have no restrictions on the matrix $\mathbf{\Phi}$

→ MLEs are often very sensitive to model specification and sample period

Literature

- NI, S. & SUN, D. (2005): Bayesian Estimates for Vector Autoregressive Models. *Journal of Business & Economic Statistics*, 23: 105–117.
- LÜTKEPOHL, H. (1993): *Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin, Heidelberg.
- BROCKWELL, P.J. & DAVIS, R.A. (1996): *Introduction to Time Series and Forecasting*. Springer Verlag, New York.
- MANSMANN, U. (2006): Skript zur Vorlesung Zeitreihenanalyse.