DSGE Methods

Identification of DSGE models

Willi Mutschler, M.Sc.

Institute of Econometrics and Economic Statistics
University of Münster
willi.mutschler@uni-muenster.de

Summer 2014

Identification

Identification problem

- Distinct parameter values do not lead to distinct probability distribution of data
- Source of identification influences findings
- Lack of identification leads to wrong conclusions from calibration and estimation

Identification of DSGE models

- concerned with two mappings
  - uniqueness of solution
    - $\hookrightarrow$ from the deep parameters to the reduced-form parameters
  - uniqueness of probability distribution
    - $\hookrightarrow$ from the solution to observable data
- Bayesian approach
  - circumvents badly shaped likelihood by using tight priors
  - comparison of prior and posterior can be misleading

Identification

- DSGE context: formal identification criteria via
  - (i) the autocovariogram (Iskrev 2010)
  - (ii) the spectral density (Komunjer and Ng 2011 & Qu and Tkachenko 2012)
  - (iii) Bayesian indicators (Koop, Pesaran, and Smith 2012)

Identification

Basic idea

- Question of uniqueness of a solution, i.e. injectivity of functions
- Formally, given an objective function $\mathfrak{f}(\theta)$ a sufficient condition for θ_0 being globally identified is given by

$$\mathfrak{f}(\theta_1) = \mathfrak{f}(\theta_0) \Rightarrow \theta_1 = \theta_0 \quad \text{for any } \theta_1 \in \Theta$$

- If this is only true in an open neighborhood of θ_0, the identification of θ_0 is local
- Population moments (mean, variance, autocovariance, spectrum) are functions of data

$\Rightarrow$ Check, whether mapping from θ to population moments is unique

- Conditions are in general only sufficient, unless data is generated from a Gaussian distribution (very common for DSGE)

Identification

Approaches to time series analysis

Time series analysis: try to explain regular behavior/patterns

Time domain approach

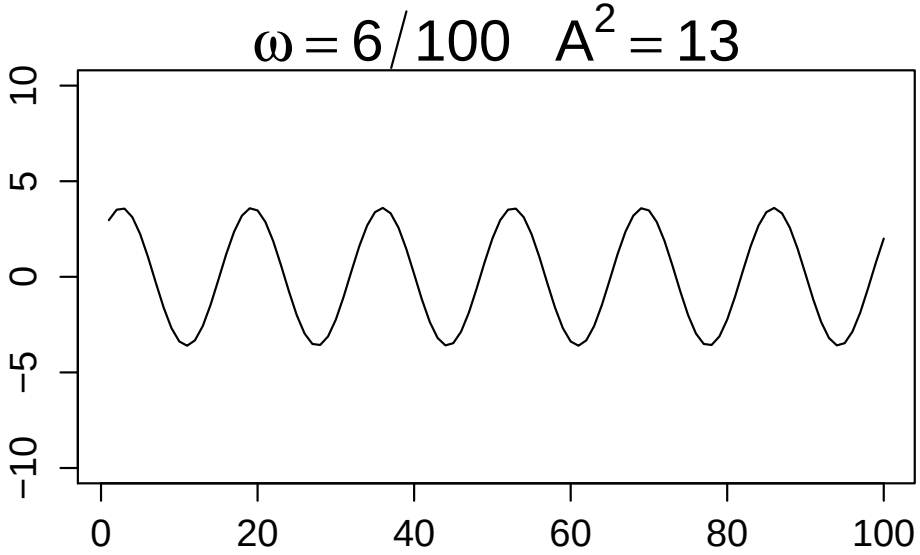
- Idea: Regression of the present on the past
- Objective: Identify bunch of parameters
- Fundamental representation via autocovariances or Wold decomposition

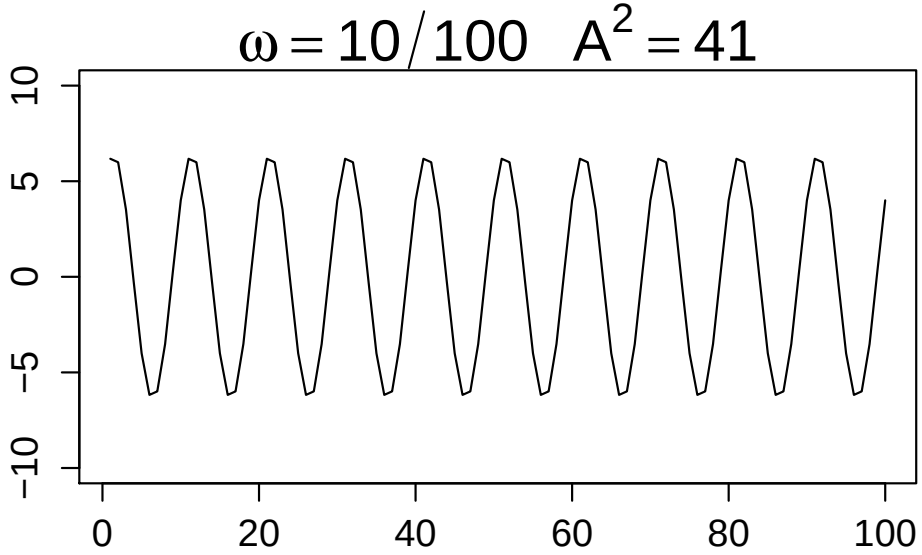
Frequency domain approach

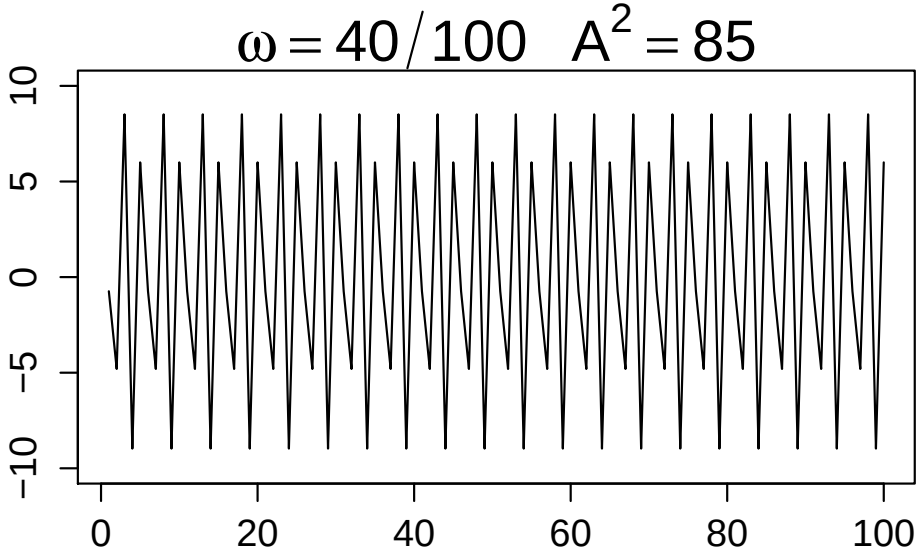
- Idea: Regression of the present on periodic sines and cosines, i.e. decomposition into regular components
- Regularity: periodic variations expressed as Fourier frequencies, driven by sines and cosines
- Objective: Identify dominant frequencies in a series
- Fundamental representation via spectral density

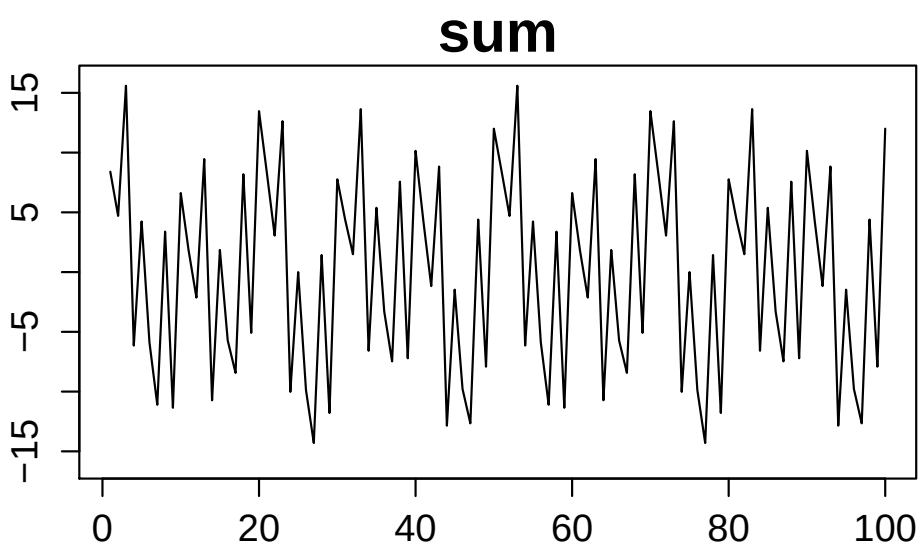
Identification

Approaches to time series analysis









Iskrev's approach

Identification

Iskrev's approach

Iskrev (2010)'s approach:

- Idea: Check whether derivative of the mean and the predicted autocovariogram of observables w.r.t deep parameters has full rank
  $\hookrightarrow$ Time domain approach
- Collect for $t = 0, 1, \ldots, T-1$ all elements in a vector

$$m(\theta, T) := \begin{pmatrix} \mu_d' & vech(\Sigma_d)' & vec(\Sigma_d(1))' & \ldots & vec(\Sigma_d(T-1))' \end{pmatrix}'$$

- If $m(\theta, q)$ is a continuously differentiable function of θ, then θ_0 is locally identifiable if the Jacobian

$$M(q) := \frac{\partial m(\theta_0, q)}{\partial \theta'}$$

has full column rank at θ_0 for $q \leq T$, i.e. equal to n_θ.

Identification

Iskrev's approach

Iskrev (2010)'s necessary condition:

- Stack all elements of the mean and the solution matrices that depend on θ into a vector τ:

$$\tau(\theta) := \begin{pmatrix} \mu_d' & vec(h_x)' & vec(g_x)' & vech(\eta_x \eta_x')' & vech(\eta_d \eta_d')' \end{pmatrix}'$$

  and consider the factorization

$$M(q) = \frac{\partial m(\theta, q)}{\partial \tau(\theta)'} \frac{\partial \tau(\theta)}{\partial \theta'}.$$

- θ_0 is locally identifiable if the rank of $J := \frac{\partial \tau(\theta_0)}{\partial \theta'}$ at θ_0 is equal to n_θ.
- This condition is, however, only necessary, because τ may be unidentifiable.

Identification

Iskrev's approach

Implementation and interpretation:

- Compute derivatives analytically or numerically
- If θ is identifiable, $M(q)$ and is likely to have full rank for q much smaller than T
- Rank deficiency: Evaluating the nullspace one can pinpoint problematic parameters

Komunjer and Ng's approach

Identification

Komunjer and Ng's approach

- Based upon results from control theory for minimal systems (observable input and output sequences):
  - minimality and invertibility are enough to characterize observational equivalence of spectral densities
- Consider minimal DSGE model, i.e. dynamics are entirely driven by the smallest possible dimension of the state vector (and shocks)
- Derive **restrictions** implied by **equivalent spectral densities** without computing any autocovariances or the spectral density

$\hookrightarrow$ *(Time and) Frequency domain approach*

Identification

Komunjer and Ng's approach

- Consider minimal DSGE model

$$x_{2,t} = \overline{x}_2 + \widetilde{\mathcal{A}}(x_{2,t-1} - \overline{x}_2) + \widetilde{\mathcal{B}}\varepsilon_t,$$
$$d_t = \overline{d} + \widetilde{\mathcal{C}}(x_{2,t-1} - \overline{x}_2) + \widetilde{\mathcal{D}}\varepsilon_t.$$

  i.e. smallest possible dimension n_{x_2} of the state vector such that dynamics are entirely driven by $x_{2,t}$ and ε_t.

  (i) Controllability: For any initial state, it is always possible to design an input sequence that puts the system in the desired final state, i.e. the matrix $\left[\widetilde{\mathcal{B}} \quad \widetilde{\mathcal{A}}\widetilde{\mathcal{B}} \quad \ldots \quad \widetilde{\mathcal{A}}^{n_{x_2}-1}\widetilde{\mathcal{B}}\right]$ has full row rank,

  (ii) Observability: Given the evolution of the input it is always possible to reconstruct the initial state by observing the evolution of the output, i.e. the matrix $\left[\widetilde{\mathcal{C}}' \quad \widetilde{\mathcal{A}}'\widetilde{\mathcal{C}}' \quad \ldots \quad \widetilde{\mathcal{A}}^{n_{x_2}-1'}\widetilde{\mathcal{C}}'\right]'$ has full column rank.

Identification

Komunjer and Ng's approach

- Since d_t is weakly stationary and ε_t is either white noise or iid:

$$d_t = \bar{d} + \widetilde{\mathcal{D}} e_t + \sum_{j=1}^{\infty} \widetilde{\mathcal{C}} \widetilde{\mathcal{A}}^{j-1} \widetilde{\mathcal{B}} \varepsilon_{t-j} = \bar{d} + \widetilde{H}_e(L^{-1}; \theta) \varepsilon_t,$$

where L is the lag-operator. For $z \in \mathbb{C}$ the transfer function (z-transform) is

$$\widetilde{H}_\varepsilon(z; \theta) := \widetilde{\mathcal{D}} + \sum_{j=1}^{\infty} \widetilde{\mathcal{C}} \widetilde{\mathcal{A}}^{j-1} \widetilde{\mathcal{B}} z^{-j} = \widetilde{\mathcal{D}} + \widetilde{\mathcal{C}} \left[I_{n_{x_2}} - \widetilde{\mathcal{A}} z \right]^{-1} \widetilde{\mathcal{B}}.$$

Identification

Komunjer and Ng's approach

- Collect all hyperparameters of the state space solution into a vector $\Lambda(\theta) := \left(vec(\widetilde{\mathcal{A}})', vec(\widetilde{\mathcal{B}})', vec(\widetilde{\mathcal{C}})', vec(\widetilde{\mathcal{D}})', vech(\Sigma_\varepsilon)' \right)'$
- For all $z \in \mathbb{C}$ the spectral density matrix of d_t is defined as

$$\Omega_d(z;\theta) := \Sigma_d + \sum_{j=1}^{\infty} \Sigma_d(j) z^{-j} + \sum_{j=1}^{\infty} \Sigma_d(-j) z^{-j} = \widetilde{H}_\varepsilon(z;\theta) \Sigma_\varepsilon(\theta) \widetilde{H}_\varepsilon(z^{-1};\theta)'.$$

- Observational equivalence of θ_0 and θ_1 is equivalent to $\forall z \in \mathbb{C}$:

$$\widetilde{H}_\varepsilon(z;\Lambda(\theta_1)) \cdot \Sigma_\varepsilon(\theta_1) \cdot \widetilde{H}_\varepsilon(z^{-1};\Lambda(\theta_1))' = \widetilde{H}_\varepsilon(z;\Lambda(\theta_0)) \cdot \Sigma_\varepsilon(\theta_0) \cdot \widetilde{H}_\varepsilon(z^{-1};\Lambda(\theta_0))'$$

  implies $\theta_0 = \theta_1$.
- Equivalent spectral densities arise if
  - (i) for given $\Sigma_\varepsilon(\theta)$, each transfer function $\widetilde{H}_\varepsilon(z;\Lambda(\theta))$ is potentially obtained from a multitude of quadruples $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \widetilde{\mathcal{C}}, \widetilde{\mathcal{D}})$,
  - (ii) there are many pairs $\widetilde{H}_\varepsilon(z;\Lambda(\theta))$ and $\Sigma_\varepsilon(\theta)$ that jointly generate the same spectral density.

Identification

Komunjer and Ng's approach

- Komunjer and Ng (2011) show that this is equivalent to the existence of a $n_{x_2} \times n_{x_2}$ similarity transformation matrix T and a $n_\varepsilon \times n_\varepsilon$ full column rank matrix $U = L_\varepsilon(\theta_0) V L_\varepsilon(\theta_1)^{-1}$ such that

$$\widetilde{\mathcal{A}}(\theta_1) = T\widetilde{\mathcal{A}}(\theta_0)T^{-1}, \quad \widetilde{\mathcal{B}}(\theta_1) = T\widetilde{\mathcal{B}}(\theta_0)U, \quad \widetilde{\mathcal{C}}(\theta_1) = \widetilde{\mathcal{C}}(\theta_0)T^{-1}$$
$$\widetilde{\mathcal{D}}(\theta_1) = \widetilde{\mathcal{D}}(\theta_0)U, \quad \Sigma_\varepsilon(\theta_1) = U^{-1}\Sigma_\varepsilon(\theta_0)U^{-1'},$$

  with L_ε being the Cholesky decomposition of $\Sigma_\varepsilon(\theta) = L_\varepsilon L_\varepsilon'$ and V a constant matrix such that $VV' = I$

$$\Rightarrow \textbf{Now we have mappings}!$$

Identification

Komunjer and Ng's approach

- Define a continuously differentiable mapping $\delta : \Theta \times \mathbb{R}^{n_{x_2}^2} \times \mathbb{R}^{n_\varepsilon^2} \to \mathbb{R}^{n_\wedge}$ as

$$\delta(\theta, T, U) := \begin{pmatrix} vec(T\widetilde{\mathcal{A}}(\theta)T^{-1}) \\ vec(T\widetilde{\mathcal{B}}(\theta)U) \\ vec(\widetilde{\mathcal{C}}(\theta)T^{-1}) \\ vec(\widetilde{\mathcal{D}}(\theta_0)U) \\ vech(U^{-1}\Sigma_\varepsilon(\theta_0)U^{-1'}) \end{pmatrix}.$$

- θ is now locally identifiable from the spectral density (or equivalently autocovariances) of d_t at a point θ_0 if and only if $\delta^S(\theta, T, U)$ is locally injective at $(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})$
- Sufficient condition: matrix of partial derivatives of $\delta^S(\theta, T, U)$ has full column rank at $(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})$

Identification

Komunjer and Ng's approach

$$
\begin{aligned}
\Delta(\theta_0) &:= \left(\frac{\partial \delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial \theta'} \quad \frac{\partial \delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial vec(T)'} \quad \frac{\partial \delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial vec(U)'} \right) \\
&= \begin{pmatrix}
\frac{\partial vec(\widetilde{\mathcal{A}})}{\partial \theta'} & \widetilde{\mathcal{A}}' \otimes I_{n_{x_2}} - I_{n_{x_2}} \otimes \widetilde{\mathcal{A}} & 0_{n_{x_2}^2 \times n_\varepsilon^2} \\
\frac{\partial vec(\widetilde{\mathcal{B}})}{\partial \theta'} & \widetilde{\mathcal{B}}' \otimes I_{n_{x_2}} & I_{n_\varepsilon} \otimes \widetilde{\mathcal{B}} \\
\frac{\partial vec(\widetilde{\mathcal{C}})}{\partial \theta'} & -I_{n_{x_2}} \otimes \widetilde{\mathcal{C}} & 0_{n_d n_{x_2} \times n_\varepsilon^2} \\
\frac{\partial vec(\widetilde{\mathcal{D}})}{\partial \theta'} & 0_{n_d n_{x_2} \times n_{x_2}^2} & I_{n_\varepsilon} \otimes \widetilde{\mathcal{D}} \\
\frac{\partial vech(\Sigma_e)}{\partial \theta'} & 0_{(n_\varepsilon(n_\varepsilon+1)/2) \times n_{x_2}^2} & -2\Xi_{n_\varepsilon}[\Sigma_e \otimes I_{n_\varepsilon}]
\end{pmatrix} \\
&=: \begin{pmatrix} \Delta_\Lambda(\theta_0) & \Delta_T(\theta_0) & \Delta_U(\theta_0) \end{pmatrix}
\end{aligned}
$$

with Ξ_{n_ε} being the left-inverse of the $n_\varepsilon^2 + n_\varepsilon(n_\varepsilon+1)/2$ duplication matrix $\mathcal{G}_{n_\varepsilon}$ for $vech(\Sigma_e)$

Order and rank condition

- Order (necessary): $n_\theta + n_{x_2}^2 + n_\varepsilon^2 \leq n_\Lambda^S := (n_{x_2} + n_d)(n_{x_2} + n_\varepsilon) + n_\varepsilon(n_\varepsilon+1)/2$
- Rank (necessary and sufficient): $rank(\Delta^s(\theta_0)) = n_\theta + n_{x_2}^2 + n_\varepsilon^2$

Identification

Komunjer and Ng's approach

Implementation and interpretation

- Analytical or numerical derivatives
- Derive minimal state vector: Check observability and controllability
- Remove the entries from all solution matrices and its derivatives corresponding to redundant state variables
- Model diagnostics via other submatrices
- Nullspace can be used to pinpoint problematic parameters

Qu and Tkachenko's approach

Identification

Qu and Tkachenko's approach

- Idea: Check whether derivative of the spectrum of observables w.r.t the deep parameters has full rank $\hookrightarrow$ Frequency domain approach
- Assume that d_t is covariance-stationary:

$$d_t = \bar{d} + \sum_{j=0}^{\infty} Dg_x h_x^j \sigma \eta_x \varepsilon_{t-j} + \eta_d \varepsilon_t = \bar{d} + H_\varepsilon(L;\theta)\varepsilon_t$$

  with $H_\varepsilon(L;\theta) = Dg_x \left(I_{n_x} - h_x L\right)^{-1} \sigma \eta_x + \eta_d$
- Using the Fourier transformation the spectral density matrix Ω_d is

$$\Omega_d(\omega,\theta) = \frac{1}{2\pi} H_\varepsilon(e^{-i\omega};\theta) \cdot \Sigma_\varepsilon(\theta) \cdot H_\varepsilon(e^{-i\omega};\theta)^*, \qquad \omega \in [-\pi;\pi],$$

  with * denoting the conjugate transpose of a complex valued matrix

Identification

Qu and Tkachenko's approach

- The test focuses on

$$\underbrace{G(\theta_0)}_{n_\theta \times n_\theta} = \int_{-\pi}^{\pi} \left(\frac{\partial vec(\Omega_d(\omega;\theta_0))'}{\partial \theta'} \right)' \left(\frac{\partial vec(\Omega(\omega;\theta_0))'}{\partial \theta'} \right) d\omega + \frac{\partial \mu_d(\theta_0)'}{\partial \theta} \frac{\partial \mu_d(\theta_0)}{\partial \theta'}.$$

- θ is locally identifiable at a point θ_0 from the mean and spectrum of d_t if and only if $G(\theta_0)$ is nonsingular, that is its rank is equal to n_θ.

Identification

Qu and Tkachenko's approach

Implementation and interpretation

- Derivative of $\Omega_d(\omega;\theta_0)$ can be calculated analytically or numerically, but numerically approximate integral
- Full rank G or $\overline{G}$: identification via spectrum or mean and spectrum
- Rank deficiency: check whether the diagonal elements of G or $\overline{G}$ are nonzero

Exercise: Identification of An and Schorfheide (2007)

Exercise: Identification of An and Schorfheide (2007)

Consider the An and Schorfheide (2007) model

1. Write a mod file for the log-linearized model.
2. Run and interpret Dynare's identification command.
3. Harder: Implement the other criteria based on ranks using numerical derivatives. Compare the results.

Bayesian identification criteria

Bayesian identification criteria

Koop, Pesaran, and Smith (2013)'s approach

Koop, Pesaran, and Smith (2013, JBES)'s approach

- Indicates weak identification using Bayesian simulation techniques (Bayesian-learning-rate-indicator)
- Rate at which the posterior precision of a given parameter gets updated with the sample size T using simulated data
- For not identified or weakly identified parameters this rate is slower than T

Bayesian identification criteria

Koop, Pesaran, and Smith (2013)'s approach

Implementation and interpretation

- Simulate data, divide into growing subsamples
- Calculate variance or precision via
  - Hessian (analytically or numerically) at posterior mode
  - MCMC variance of posterior distribution
- Precision divided by sample size will go to
  - zero for weakly identified parameters
  - a constant for identified parameters

Comparison

Comparison

- Methods differ due to different perspectives and assumptions:
  - Iskrev: time domain, very general assumptions (moments), compute derivative of autocovariogram
  - Qu/Tkachenko: frequency domain, very general assumptions (VMA representation), compute derivative of spectrum
  - Komunjer/Ng: (time and) frequency domain, strictest assumptions (minimality and left-invertibility), conditions without computing autocovariances or spectral density
  - KPS: Bayesian simulation approach
- Computational issues and numerical errors
  - If feasible: Use analytical derivatives rather than numerical
  - Isk: lag length, QT: subintervals for integral, KN: Find minimal representation, KPS: Filtering, Speed
  - For all: numerical instability of solution algorithm, size of matrices
- Methods include different model diagnostics