Identification Analysis of DSGE Model Parameters with Dynare

Willi Mutschler

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Dynare Team

Motivation

Motivation

Estimation of DSGE models has rapidly progressed

Parameter identification

- is a r*egularity condition*
- can and should be assessed *before* taking your model to data
- is a *model property*

Dynare's Identification Toolbox

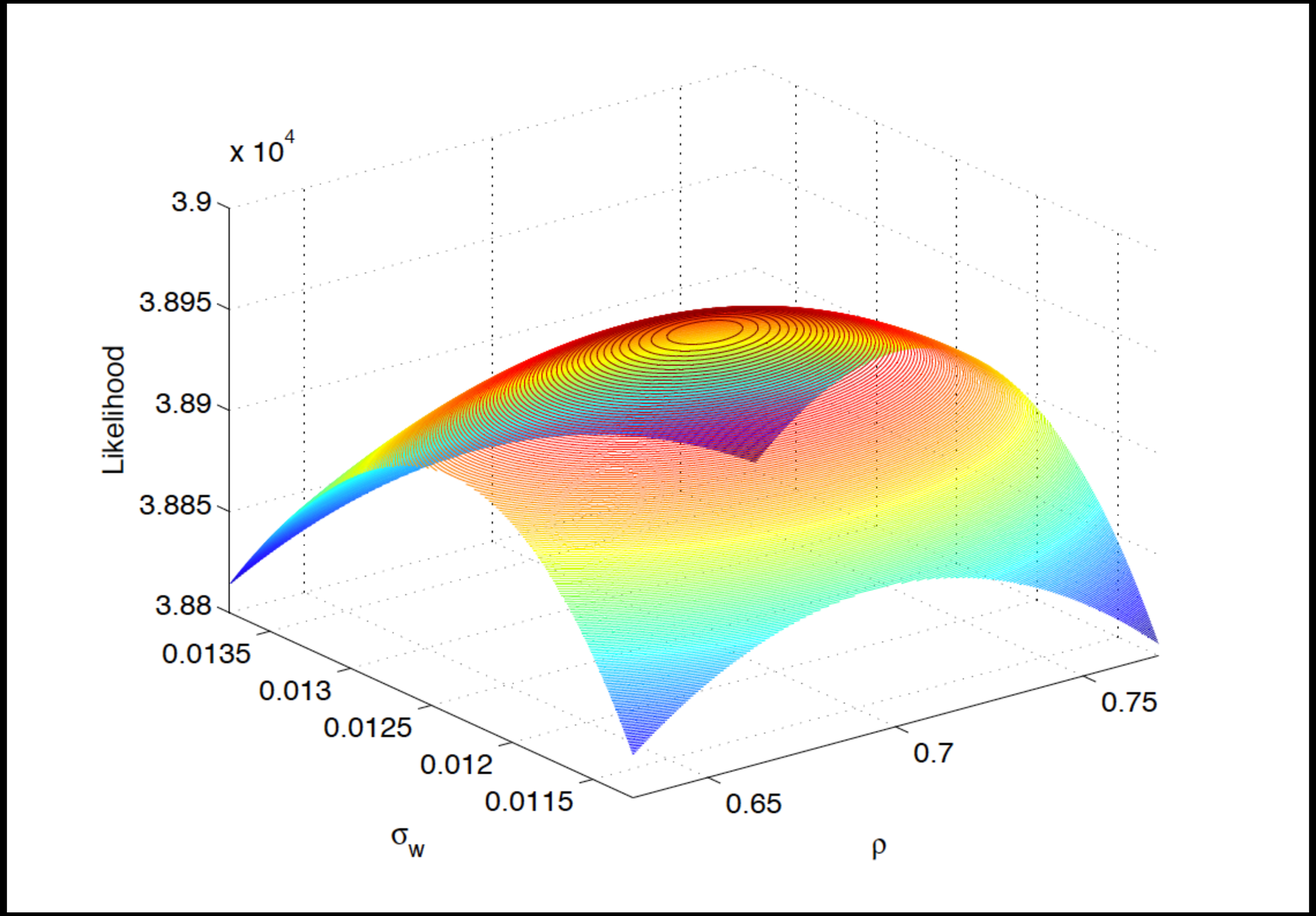
Based on well-established diagnostics and indicators to detect

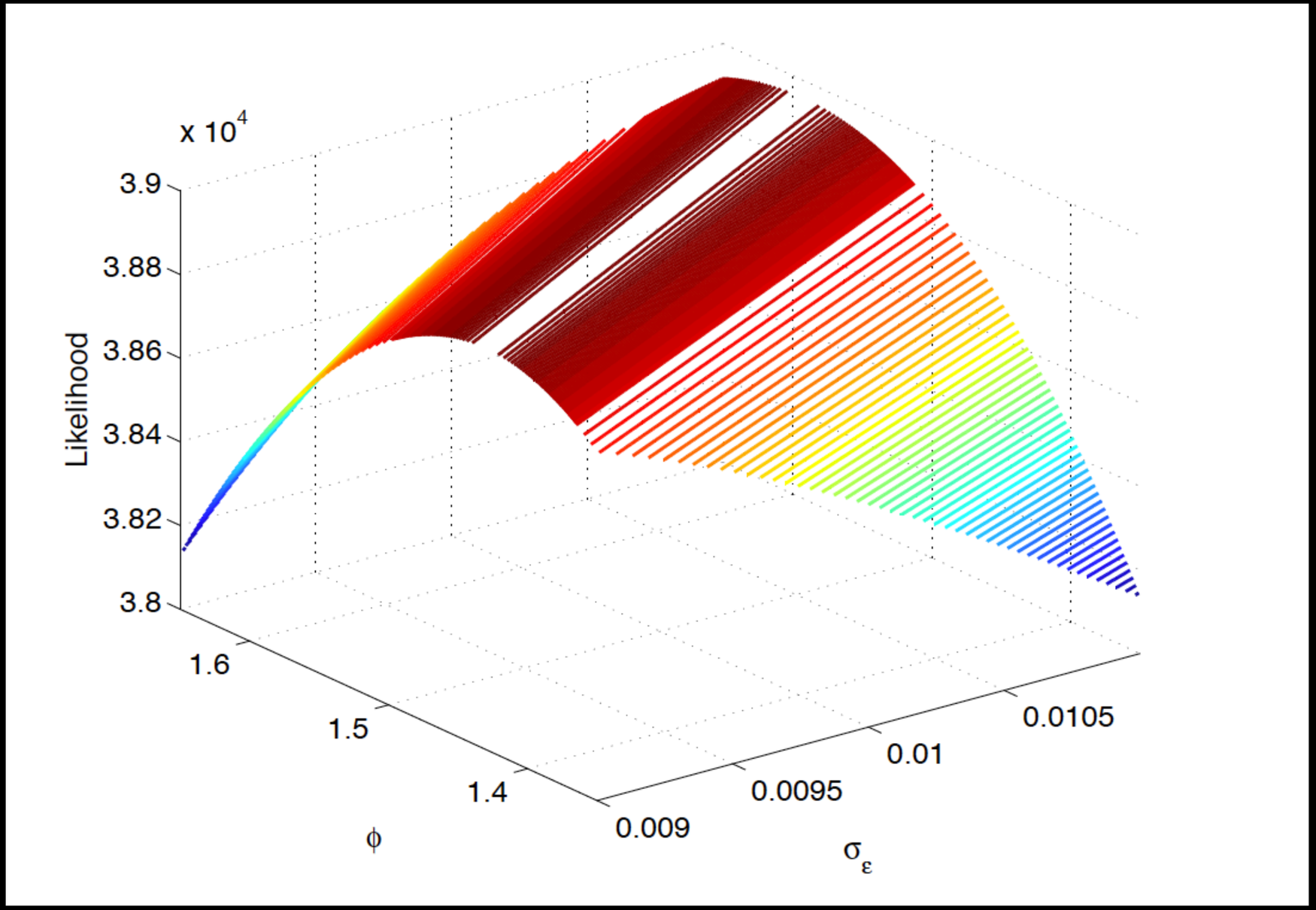
- theoretical identification failures
- weak identification

Feature requests and contributions are very welcome!

Example 1

Likelihood shape

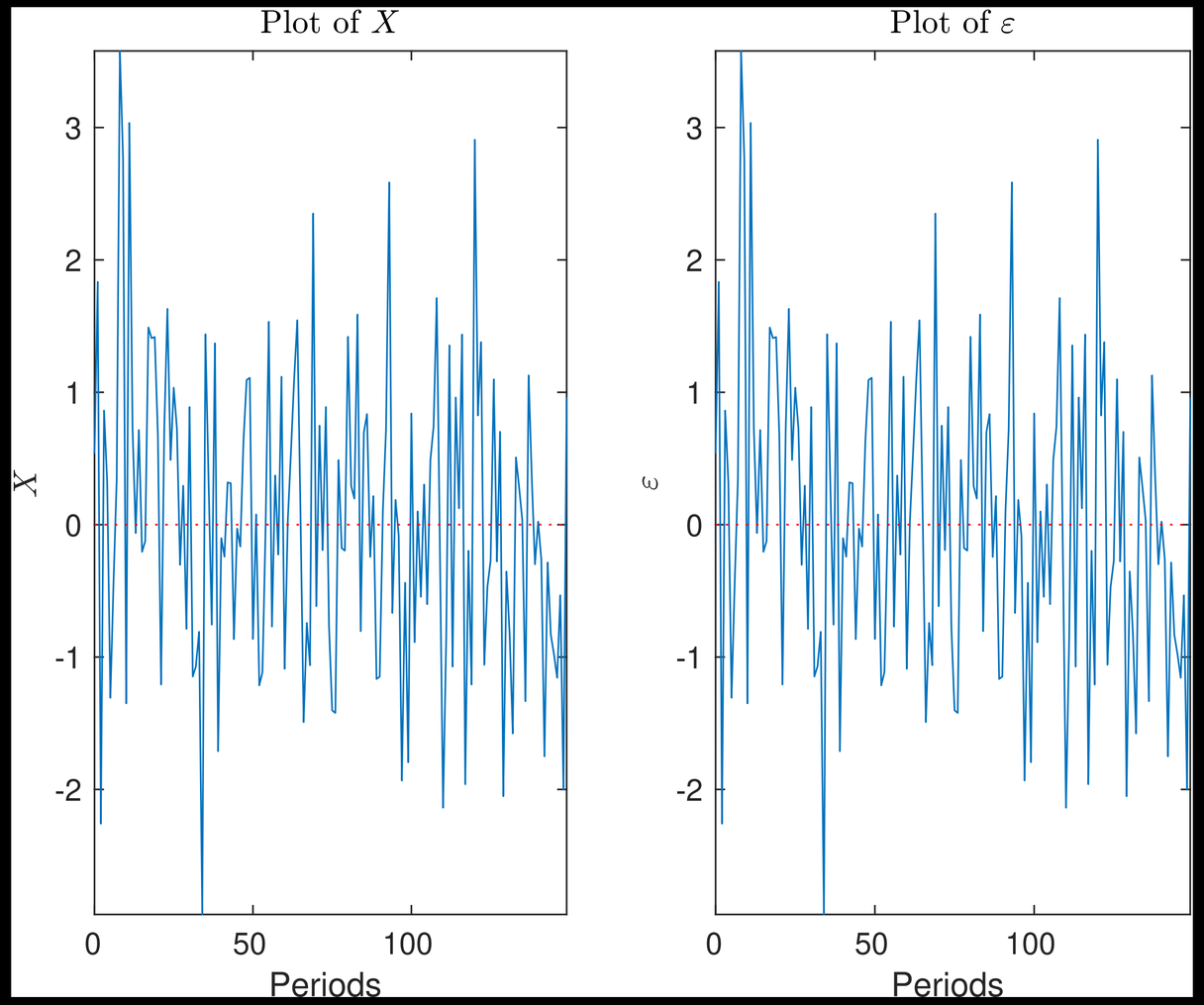

x 10^4
3.9
3.895
3.89
3.885
3.88
Likelihood
0.0135
0.013
0.0125
0.012
0.0115
σw
0.65
0.7
0.75
ρ


x 10^4
3.9
3.88
3.86
3.84
3.82
3.8
Likelihood
1.6
1.5
1.4
ϕ
0.009
0.0095
0.01
0.0105
σε

Example 2

ARMA(1,1)

$$x_t - \phi_1 x_{t-1} = \varepsilon_t - \phi_2 \varepsilon_{t-1}, \text{ with } \varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$$



arma_common.inc
arma_plots.mod

ARMA(1,1)

ARMA(1,1) autocovariance function (define $\gamma_j = E[x_t x_{t-j}]$):

$$\gamma_0 = \frac{(1 + \phi_2^2 - 2\phi_1\phi_2)\sigma^2}{1 - \phi_1^2}, \quad \gamma_1 = \frac{(\phi_1 - \phi_2)(1 - \phi_1\phi_2)\sigma^2}{1 - \phi_1^2}, \quad \gamma_h = \phi_1\gamma_{h-1}$$

Case 1: equal coefficients ($\phi_1 = \phi_2$) autocovariance function:

$$\gamma_0 = \sigma^2, \gamma_h = 0$$

Case 2: White Noise ($\phi_1 = \phi_2 = 0$) autocovariance function:

$$\gamma_0 = \sigma^2, \gamma_h = 0$$

ARMA(1,1)

Observational equivalence between cases 1 and 2

```
SPECTRUM (QU AND TKACHENKO, 2012):
    !!!WARNING!!!
    The rank of Gbar (Jacobian of mean and spectrum) is deficient!

    [THETA_2,THETA_1] are PAIRWISE collinear!

MOMENTS (ISKREV, 2010):
    !!!WARNING!!!
    The rank of J (Jacobian of first two moments) is deficient!

    [THETA_2,THETA_1] are PAIRWISE collinear!
```

arma_identif.mod

Example 3

Forward-looking DSGE model

Forward-looking DSGE model

Taylor-rule: $r_t = \psi \pi_t + \varepsilon_t^M$

Dynamic IS: $x_t = E_t x_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1}) + \varepsilon_t^D$

NK-Phillips curve: $\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + \varepsilon_t^S$

Forward-looking model

$$\underbrace{\begin{pmatrix} 1 & 0 & -\psi \\ \frac{1}{\tau} & 1 & 0 \\ 0 & -\kappa & 1 \end{pmatrix}}_{A_0} \underbrace{\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\tau} \\ 0 & 0 & \beta \end{pmatrix}}_{A_1} \underbrace{\begin{pmatrix} E_t r_{t+1} \\ E_t x_{t+1} \\ E_t \pi_{t+1} \end{pmatrix}}_{E_t y_{t+1}} + \underbrace{\begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}}_{\varepsilon_t}$$

Stationary solution implies Eigenvalues of $A_0^{-1}A_1$ lie within unit circle

$$y_t = A_0^{-1}A_1 E_t y_{t+1} + A_0^{-1}\varepsilon_t = \sum_{j=0}^{\infty} (A_0^{-1}A_1)^j A_0^{-1} E_t \varepsilon_{t+j} = A_0^{-1}\varepsilon_t$$

`forward_looking_common.inc`

Forward-looking model

Stationary solution implies Eigenvalues of $A_0^{-1}A_1$ lie within unit circle

$$y_t = A_0^{-1}A_1 E_t y_{t+1} + A_0^{-1}\varepsilon_t = \sum_{j=0}^{\infty} (A_0^{-1}A_1)^j A_0^{-1} E_t \varepsilon_{t+j} = A_0^{-1}\varepsilon_t$$

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1} \begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Forward-looking model

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

- Parameter space is not *variation free* (bounds on parameters to ensure Eigenvalues of $A_0^{-1}A_1$ inside unit circle)
- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)

- κ is product of several other structural parameters (Calvo or Rotemberg)
- Identification depends on observables, e.g. when observing only x_t
  - τ and κ are pairwise colinear
  - κ and ψ are pairwise colinear
  - we would need to fix e.g. ψ

Forward-looking model

```
varobs x;
```

```
REDUCED-FORM:
    !!!WARNING!!!
    The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

    BETA is not identified!

SPECTRUM (QU AND TKACHENKO, 2012):
    !!!WARNING!!!
    The rank of Gbar (Jacobian of mean and spectrum) is deficient!

    BETA is not identified!
    [TAU,PSI] are PAIRWISE collinear!
    [KAPPA,PSI] are PAIRWISE collinear!
```

forward_looking_varobs_x.mod

Forward-looking model

```
varobs r x p;
```

```
REDUCED-FORM:
    !!!WARNING!!!
    The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

    BETA is not identified!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
    !!!WARNING!!!
    The rank of Deltabar (Jacobian of steady state and minimal system) is deficient!

    BETA is not identified!

SPECTRUM (QU AND TKACHENKO, 2012):
    !!!WARNING!!!
    The rank of Gbar (Jacobian of mean and spectrum) is deficient!

    BETA is not identified!

MOMENTS (ISKREV, 2010):
    !!!WARNING!!!
    The rank of J (Jacobian of first two moments) is deficient!

    BETA is not identified!
```

forward_looking_varobs_all.mod

Which observables?

Observing all variables would be best, but unrealistic as we face typical trade-offs:

- what can you actually observe in data?
- how to match model variables to data?
- some estimation techniques require as many observables as shocks

Some guidance:

- Different combinations of observables influence identification
- dynare_sensitvitiy helps get intuition and insights which parameters are most important for which variables
- Brute-force: try out all combinations and take the one that works best

Example 4

Investment Adjustment Costs

Kim (2003)

RBC model: $U_t = \ln(C_t),\ Y_t = A_t K_{t-1}^{\alpha},\ \ln(A_t) = \rho_A \ln(A_{t-1}) + \varepsilon_t^A$

Two types of investment adjustment costs

$$Y_t^d = \left[(1 - SAV) \left(\frac{C_t}{1 - SAV} \right)^{1+\theta} + SAV \left(\frac{I_t}{SAV} \right)^{1+\theta} \right]^{\frac{1}{1+\theta}}, \text{ with } SAV = \frac{\bar{I}}{\bar{Y}^d}$$

$$K_t = \left[(1 - \delta) {K_{t-1}}^{1-\kappa} + \delta \left(\frac{I_t}{\delta} \right)^{1-\kappa} \right]^{\frac{1}{1-\kappa}}$$

investadjust_common.inc

Kim (2003): Different parametrizations

$\kappa = 1.4$ and $\theta = 0$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.3724	0.1387
q	1.0000	0.7295	0.5322
a	1.0000	0.6928	0.4800

investadjust_nomulticosts.mod

$\kappa = 2$ and $\theta = 1.5$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.4108	0.1688
q	1.0000	0.8094	0.6551
a	1.0000	0.6928	0.4800

investadjust_baseline.mod

$\kappa = 0$ and $\theta = -3.5$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.2832	0.0802
q	1.0000	0.5448	0.2968
a	1.0000	0.6928	0.4800

investadjust_nointertempcosts.mod

Kim (2003): varobs c iv

(θ, κ) and κ are jointly non-identifiable as long as $\frac{\theta + \kappa}{1 + \theta}$ stays the same value

```
REDUCED-FORM:
    !!!WARNING!!!
    The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

    RA is not identified!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
    !!!WARNING!!!
    The rank of Deltabar (Jacobian of steady state and minimal system) is deficient!

    RA is not identified!
    [KAPPA,THETA] are PAIRWISE collinear!

SPECTRUM (QU AND TKACHENKO, 2012):
    !!!WARNING!!!
    The rank of Gbar (Jacobian of mean and spectrum) is deficient!

    RA is not identified!
    [KAPPA,THETA] are PAIRWISE collinear!

MOMENTS (ISKREV, 2010):
    !!!WARNING!!!
    The rank of J (Jacobian of first two moments) is deficient!

    RA is not identified!
    [KAPPA,THETA] are PAIRWISE collinear!
```

investadjust_identif.mod

Kim (2003): varobs c iv lam

Observing lagrange multipliers λ or Tobin's Q provides means to identify (θ, κ) separately

What about r^A?

```
REDUCED-FORM:
    !!!WARNING!!!
    The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

    RA is not identified!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
    !!!WARNING!!!
    The rank of Deltabar (Jacobian of steady state and minimal system) is deficient!

    RA is not identified!

SPECTRUM (QU AND TKACHENKO, 2012):
    !!!WARNING!!!
    The rank of Gbar (Jacobian of mean and spectrum) is deficient!

    RA is not identified!

MOMENTS (ISKREV, 2010):
    !!!WARNING!!!
    The rank of J (Jacobian of first two moments) is deficient!

    RA is not identified!
```

investadjust_identif_lam.mod

Kim (2003): What about r^A

typical failure: $\beta = \dfrac{1}{1 + r^A/400}$

```
parameters
BETA    ${\beta}$     (long_name='discount factor')
RA      ${r_{A}}$     (long_name='annual steady-state real interest rate (defines discount factor)')
```

```
RA          = 2;              % annual nominal interest rate
BETA        = 1/(1+RA/400); % discount factor
```

investadjust_identif.mod

Kim (2003): Insights

- Different parameters yield same theoretical moments for y, c, iv, rk and a;
  ↪ therefore also observational equivalent model dynamics (e.g. IRFs)

- At first-order, perturbation solution is linear Gaussian state space
  ↪ First two moments characterize whole distribution (higher-order moments, IRFs, variance decomposition, etc)

- Even with an infinite sample size, no way to uniquely recover true model structure

- Observing lam or Q identifies model parameters, but unrealistic

Parameter Identification in Theory and Practice

Identification Problem in Theory

Observational Equivalence: structural models with potentially different economic interpretations may be indistinguishable

Under-identification: objective function is independent of certain structural parameters, e.g. because they disappear from rational expectations solution

Partial identification: with two or more structural parameters entering objective function only proportionally, making them separately unrecoverable

Weak identification: parameter theoretically identified, but curvature may be small in certain regions of the parameter space

Identification Problem in Practice

Lack of identification leads to wrong conclusions from calibration, estimation and inference

Many caveats due to identifiability issues and/or an unfortunate choice of observables

- difficult to maximize likelihood/posterior or minimize some (moment) objective function
- estimators often lie on the boundary of theoretically admissible space
- Gaussian asymptotic theory yields poor approximations

Weak identification is likely a more serious concern for applied researchers

“Unidentifiability causes no real difficulties in the Bayesian approach”

Lindley (1971, p. 46)

Is there a systematic way to detect such issues?

Theoretical Lack of Identification

Distinct parameter values do not lead to distinct objective functions of data

Global vs. local identifiability

Definitions

$\theta \in \Theta$: vector of model parameters

Θ: admissable parameter space for unique and stable solution

Y_T: matrix of observables with sample size T

$p(\theta; Y_T)$: objective function generated by a DSGE model

Global and local identification (Rotemberg, 1971)

A point $\theta_0 \in \Theta$ is said to be *globally identified* if for all Y_T:

$$p(\theta_0; Y_T) = p(\theta_1; Y_T) \text{ implies } \theta_0 = \theta_1$$

If this is true only for values $\tilde{\theta}$ in an open neighborhood of θ_0, then θ_0 is said to be *locally identified*.

Strength of Identification

How much information can be extracted from a specific Y_T to estimate model parameters precisely?

More formally: What happens to the precision of estimates with a growing sample size.

Literature

Literature: Global Identification

- Minimizing Kullback-Leibler discrepancy (Qu and Tkachenko, 2017)
- Exploiting link between observationally equivalent state space representations and model solution constraints (Kocicecki & Kolasa, 2018)

Literature: Local Identification

For linearized Gaussian DSGE models

✓ Moments (Iskrev, 2010; Ratto and Iskrev, 2011)

✓ Spectral density (Qu and Tkachenko, 2012)

✓ Control theory for minimal systems (Komunjer and Ng, 2011)

Literature: Local Identification

For non-linear and / or non-Gaussian DSGE models

✓ Cumulants and Polyspectra (Mutschler, 2015)

✓ Adding mean restrictions to minimal first-order system (Mutschler, 2014)

Literature: Weak Identification

For linearized (and non-linear) DSGE models

✓ Asymptotic information matrix (Ratto and Iskrev, 2011; Andrle, 2010)

✳ Bayesian indicators (Koop, Pesaran and Smith, 2013)

◉ Indirect inference on VAR approximation (Le, Meenagh, Minford and Wickens, 2017)

◉ Score test on Gaussian likelihood (Qu, 2014)

Identification in linearized Gaussian DSGE models

Linear Gaussian state space system

Dynare's first-order perturbation solution for endogenous variables z_t given state variables x_t and shocks u_t:

$$z_t - \bar{z} = ghx \cdot (x_{t-1} - \bar{x}) + ghu \cdot u_t$$

Rewrite to more familiar state-space form:

$$\begin{aligned} x_t - \bar{x} &= A(\theta)(x_{t-1} - \bar{x}) + B(\theta)u_t \quad \text{[STATES]} \\ y_t - \bar{y} &= C(\theta)(x_{t-1} - \bar{x}) + D(\theta)u_t \quad \text{[VAROBS]} \end{aligned}$$

where $u_t \sim N(0,\Sigma_u)$

Linear state space system

As $u_t \sim \mathcal{N}(0, \Sigma_u)$, the unconditional first and second moments are

$$E[y_t] \equiv \mu_y = \bar{y}, \quad E(x_t) \equiv \mu_x = \bar{x}$$

$$E[(y_t - \bar{y})(y_t - \bar{y}')] \equiv \Sigma_y(0) = C\Sigma_x(0)C' + D\Sigma_u D'$$

$$E[(x_t - \bar{x})(x_t - \bar{x})'] \equiv \Sigma_x(0) = A\Sigma_x(0)A' + B\Sigma_u B'$$

where $\Sigma_x(0)$ is fixed point of Lyapunov equation

Linear state space system

From this one can compute the

- theoretical autocovariogram $\Sigma_x(j)$ and $\Sigma_y(j)$
- theoretical spectral density $\mathcal{S}_{2,x}(\omega)$ and $\mathcal{S}_{2,y}(\omega)$ for $\omega \in [-\pi; \pi]$
- minimal state space system with minimal state variables $\tilde{x}_t$

Some remarks

In general, only observing y_t would be insufficient to fully characterize the distribution of state variables x_t and all endogenous variables y_t

Fortunately, our model implies restrictions through θ

Mapping from structural parameters θ to $(A(\theta), B(\theta), C(\theta), D(\theta))$ is highly nonlinear and only implicitly available

Diagnostics based on moments

Moments (Iskrev, 2010)

Proposition: Suppose that

$$m(q) := \left[\mu_y', vech(\Sigma_y(0))', vec(\Sigma_y(1)), \ldots, vec(\Sigma_y(q))'\right]$$

is a continuously differentiable function of $\theta \in \Theta$. Let $\theta_0 \in \Theta$ be a regular point. θ is then locally identifiable at a point θ_0 from the first two moments of Y_T, if the Jacobian matrix

$$J = \frac{\partial m(q)}{\partial \theta'}$$

has full column rank at θ_0 for $q \leq T$.

This condition is both necessary and sufficient when $q = T$ if u_t is Gaussian.

Intuition

Unique (injective) mapping from structural parameters to population moments

Check injectivity by looking at rank of Jacobian matrix

- *Order condition*: At least as many moments as parameters
- *Rank condition*: Check whether Jacobian of *theoretical mean, variance and autocovariogram* of observables w.r.t structural parameters has full rank

Very helpful in detecting observational equivalence (columns of zeros and linear dependence between parameters)

Checking global identification is hard, but one can check local identification for relevant parameter range

Decomposing Jacobian of Moments

$$J = \frac{\partial m(T)}{\partial \tau'} \frac{\partial \tau}{\partial \theta'}$$

Reduced-form parameters: $\tau := \left[\bar{z}', vec(ghx)', vech(ghu \cdot \Sigma_u \cdot ghu')'\right]'$

$\frac{\partial \tau}{\partial \theta'}$: sensitivity of model solution parts τ to changes in model parameters θ

$\frac{\partial m(T)}{\partial \tau'}$: sensitivity of observed data moments to changes in model solution

Jacobian of Reduced-Form

Corollary: The point θ_0 is locally identifiable only if $\frac{\partial \tau}{\partial \theta'}$ at θ_0 has full rank.

- Necessary condition as parameters only affect distribution of observables through their effect on model solution
- It is not sufficient unless all variables are observed

Diagnostics based on spectrum

Spectrum (Qu and Tkachenko, 2012)

Proposition: Suppose that the spectral density $\mathcal{S}_{2,y}$ is continuous in ($\omega \in [-\pi; \pi]$ and continuous and differentiable in $\theta \in \Theta$. Let

$$\bar{G} = \left(\frac{\partial \mu_y}{\partial \theta'}\right)' \left(\frac{\partial \mu_y}{\partial \theta'}\right) + \int_{-\pi}^{\pi} \left(\frac{\partial \mathcal{S}_{2,y}}{\partial \theta'}\right)^* \left(\frac{\partial \mathcal{S}_{2,y}}{\partial \theta'}\right) \mathrm{d}\omega$$

and $\theta_0 \in \Theta$ be a regular point. Furthermore, assume there is an open neighborhood of θ_0 in which $\bar{G}$ has a constant rank. Then θ is locally identifiable at a point θ_0 from the mean and spectrum of Y_T, if and only if $\bar{G}$ is non-singular and equal to the number of parameters.

Intuition

Unique (injective) mapping from structural parameters to population mean and spectral density

Check injectivity by looking at rank of Jacobian matrix

- *Rank condition:* Jacobian of *theoretical mean and spectrum* of observables w.r.t structural parameters has full rank

Very helpful in detecting observational equivalence (columns of zeros and linear dependence between parameters or try all combinations of sets of parameters);

Gram matrix structure numerically facilitates rank computations, no order condition required

Checking global identification is hard, but one can check local identification for relevant parameter range

Diagnostics based on control theory for minimal systems

Minimal State Space System

Dynamics are entirely driven by the smallest possible dimension of the state vector (and shocks)

Definition of minimality:

- *Controllability*: For any initial state, it is always possible to design an input sequence that puts the system in the desired final state

- *Observability*: Given the evolution of the input it is always possible to reconstruct the initial state by observing the evolution of the output

Solution in Dynare is by default not based on the minimal state representation (but there is a function `get_minimal_state_representation.m`)

Numerical procedures (pole-zero cancellation) do not necessarily output minimal states with economic meaning

Minimal System (Komunjer and Ng, 2011)

Proposition: Consider the minimal linearized DSGE model solution. Assume that the vector containing the minimal solution matrices

$$\Lambda := \left(vec(\tilde{A})', vec(\tilde{B})', vec(\tilde{C})', vec(\tilde{D})', vech(\tilde{\Sigma}_u)'\right)'$$

is continuously differentiable on Θ. Two triples $(\theta_0, I_{\tilde{x}}, I_u)$ and (θ_1, T, U) are observationally equivalent if

$$\tilde{A}(\theta_1) = T\tilde{A}(\theta_0)T^{-1}, \tilde{B}(\theta_1) = T\tilde{B}(\theta_0)U^{-1}, \tilde{C}(\theta_1) = \tilde{C}(\theta_0)T^{-1}, \Sigma_u(\theta_1) = U^{-1}\Sigma_u(\theta_0)U^{-1}$$

with similarity transformation matrices T and U being full rank matrices.

Minimal System (Komunjer and Ng, 2011)

Proposition (continued): Let $\theta_0 \in \Theta$ be a regular point, then θ is locally identifiable at a point θ_0 from the mean, autocovariances, and spectrum of Y_T if and only if:

$$\bar{\Delta} := \begin{pmatrix} \frac{\partial \mu_y}{\partial \theta'} & 0 & 0 \\ \frac{\partial vec(\widetilde{A})}{\partial \theta'} & \widetilde{A}' \otimes I - I \otimes \widetilde{A} & 0 \\ \frac{\partial vec(\widetilde{B})}{\partial \theta'} & \widetilde{B}' \otimes I & I \otimes \widetilde{B} \\ \frac{\partial vec(\widetilde{C})}{\partial \theta'} & -I \otimes \widetilde{C} & 0 \\ \frac{\partial vec(\widetilde{D})}{\partial \theta'} & 0 & I \otimes \widetilde{D} \\ \frac{\partial vec(\Sigma_u)}{\partial \theta'} & 0 & -2[\Sigma_u \otimes I] \end{pmatrix}$$

has full column rank.

Intuition

Based upon identification results from control theory for minimal systems

Derive restrictions implied by equivalent model dynamics without computing any autocovariances or the spectral density

- for given size of shocks Σ_u, each transfer function is potentially obtained from a multitude of quadruples (A, B, C, D)
- many pairs of transfer functions and size of shocks Σ_u that jointly generate the same spectral density

Rank condition: Check injectivity of restrictions by computing rank of Jacobian matrix

Order condition and some other useful diagnostics for IRFs

Dynare Implementation

`identification;`

Triggers the local identification tests and has two modes of operation:

- Point identification check (default)
- Monte Carlo mode (e.g. `prior_mc=1000`)

By default all model parameters and all stderr parameters are checked

Parameters can be selected with *estimated_params* block either with

- initial values
- prior information

Warnings

Use best practices in your mod file!

Dynare's symbolic preprocessor interprets and implements the model definitions as expressed in the Dynare file

It will not reflect all parameter definitions which may be hidden e.g. in a custom `_steadystate.m` file (use `steady_state_model` instead)

Try to avoid declaring auxiliary parameters, but use the '#' syntax in the `model` block to declare endogenous parameters

Tracking singularities

When Jacobians are rank deficient, code tries to diagnose subset of parameters responsible

- ranks are computed using the singular value decomposition
- for columns of zeros in Jacobian, associated parameter is printed on command window
- compute pairwise- and multi-correlation coefficients for each column of Jacobian: if there are parameters with correlation coefficients equal to unity, these are printed on the command window

Alternatively, a brute-force approach to check rank conditions for all possible combinations can be triggered (`checks_via_subsets=1`)

Example:
Point vs Monte Carlo

`investadjust_identif.mod`

`investadjust_identif_mc.mod`

Computational remarks

```
Note that differences in the criteria could be due to numerical settings,
numerical errors or the method used to find problematic parameter sets.
Settings:
    Derivation mode for Jacobians:                          Analytic using sylvester equations
    Method to find problematic parameters:                  Nullspace and multicorrelation coefficients
    Normalize Jacobians:                                    Yes
    Tolerance level for rank computations:                  robust
    Tolerance level for selecting nonzero columns:          1e-08
    Tolerance level for selecting nonzero singular values: 1e-03
```

Many options you can change, see the manual

Key issue: distinguish near linear dependence (weak identification) form true perfect collinearity

Whenever possible the code uses closed-form expressions to compute the Jacobians analytically
↪ `get_perturbation_params_derivs.m`

Errors of numerical differentiations: the rank test for singularity is much more sensitive to the significance threshold set the by user in checking the rank when numerical derivatives are used

Identification Strength

Weak identification

Even though all parameters are locally identifiable (enter objective function separately and it has a unique extremum), its curvature may be small in certain regions of the parameter space (especially in small samples)

Diagnostics are based on *precision* of parameter estimates computed via

- Inverse of (asymptotic) Fisher Information Matrix (Andrle, 2010; Ratto and Iskrev, 2011)

- Bayesian Learning Rate Indicator for growing sample sizes (Koop, Pesaran, Smith, 2013)

Identification Strength via Information Matrix

The Information Matrix Revisited

Precision of parameter estimates at the mode is given (asymptotically) by the inverse of the *Fisher Information Matrix*

$$\mathcal{I}(\theta) = E\left[\left(\frac{\partial \log(p(y\,|\,\theta))}{\partial \theta'}\right)'\left(\frac{\partial \log(p(y\,|\,\theta))}{\partial \theta'}\right)\right]$$

Typically, non-singularity of this matrix is sufficient for local identification

The Information Matrix Revisited

Decompose *Fisher Information Matrix*

$$\mathcal{I}(\theta) = \Delta^{\frac{1}{2}} \tilde{\mathcal{I}}(\theta) \Delta^{\frac{1}{2}}$$

- variance matrix: $\Delta = diag(\mathcal{I}(\theta))$
- correlation matrix: $\tilde{\mathcal{I}}(\theta)$

Identification Strength Intuition

1. The likelihood does not change when parameter θ_i changes:

$$\frac{\partial \log(p(y \mid \theta))}{\partial \theta_i} = 0 \Leftrightarrow \Delta_i = 0$$

2. The effect on the likelihood is offset due to perfect correlation:

$$\rho_i \equiv corr\left(\frac{\partial \log(p(y \mid \theta))}{\partial \theta_i}, \frac{\partial \log(p(y \mid \theta))}{\partial \theta_{-i}}\right) = 1$$

↪ This suggests that we can use this curvature information also for *weak identification:*

$$\Delta_i \approx 0 \text{ and/or } \rho_i \approx 1$$

Identification Strength Formally

If $\mathcal{I}(\theta)$ is not singular, Cramer-Rao bound of the uncertainty of estimated θ_i

$$\text{std}(\theta_i) \geq \sqrt{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}$$

Identification strength can therefore be measured using (Andrle, 2010):

$$s_i(\theta_i) \equiv \text{std}(\theta_i)^{-1} = \frac{1}{\sqrt{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{\Delta_i(1-\rho_i^2)}$$

As this precision measure only uses the population objective function (expected log-likelihood), it is an a-priori measure

At the same time, it can be interpreted as the lower bound on the estimation uncertainty in an unbiased finite sample estimator

Identification Strength in Dynare

Problem: measure is percentage change in likelihood for a unit change in parameter → not unit free

Solution: multiply by the value of θ_i to get a re-normalized version of the *curvature*

$$s_i^{dyn}(\theta_i) \equiv \theta_i \times s_i(\theta_i) = \sqrt{\frac{\theta_i^2}{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{\theta_i^2 \Delta_i} \times \sqrt{(1-\rho_i^2)} \text{ [STRENGTH]}$$

The sensitivity component contained in this measure (as opposed to the correlation component):

$$\Delta_i^{dyn} = \sqrt{\theta_i^2 \Delta_i} = \sqrt{\theta_i^2 \mathcal{I}(\theta)_{(i,i)}} \text{ [SENSITIVITY]}$$

Alternative Normalization

Identification is often checked over the prior region, e.g. at the prior mean

Because some parameters may have a prior mean of zero, alternative normalization using the *prior standard deviation* $\sigma(\theta_i)$

$$s_i^{prior} = \sigma(\theta_i) / \sqrt{(\mathcal{I}(\theta)^{-1})_{(i,i)}} \quad \text{[STRENGTH]}$$

$$\Delta_i^{prior} = \sigma(\theta_i) \cdot \sqrt{\mathcal{I}(\theta)_{(i,i)}} \quad \text{[SENSITIVITY]}$$

Identification Strength Plots

The identification toolbox shows, after the check of rank conditions, the *plots of the strength of identification and of the sensitivity component* for all estimated parameters

Example: `invest_adjust_identif.mod`

Numerical Remark: Computing $\mathcal{I}(\theta)$

Asymptotic Information Matrix

- Fischer information matrix $\mathcal{I}(\theta)$ can be computed analytically; which also provides an estimate of the Cramer-Rao lower bound of uncertainty of θ

Moment Information Matrix

- perform (and repeat) stochastic simulations and compute sample moments of observed variables
- take covariance matrix of (first and second) simulated moments
- moment information matrix: $\mathcal{I}(\theta \,|\, m_T) = J \cdot \Sigma(m_T)^{-1} \cdot J'$
- does NOT provide an estimate of the Cramer-Rao lower bound
- triggered with stochastic singularity and whenever the Asymptotic Information Matrix cannot be computed

Example:
Identification Strength Plots

`investadjust_identif_mc.mod`

Identification Strength via Bayesian Learning Rate Indicator

Koop, Pesaran and Smith (2013)

Basic idea: strength of identification becomes better as more data becomes available, i.e. parameters are estimated more precisely

Suppose θ_2 is identified, whereas θ_1 is weakly identified such that the rank of the reduced-form parameters depends on the sample size T. For growing T the posterior precision of

- θ_1 divided by the sample size will go to **zero**.
- θ_2 divided by the sample size will go to a **constant**

↪ *Bayesian simulation approach*

Koop, Pesaran and Smith (2013)

Not yet implemented as an option in the toolbox, but actually very easy to do:

- Simulate large dataset, estimate with Bayesian MCMC using first e.g. 100, 300, 900, 2700... observations
- Look at either average posterior precisions or convergence ratios

Slice with double rotation (50x200x2500 draws) works very well as it requires almost no fine-tuning (and no mode-finding step)

Use parallel option to get more chains

Example: Kim (2003)

Average Posterior Precisions

T	α	R^A	δ	ρ_A	σ_A	θ	κ
100	5.02419	0.16006	68.06352	2.35966	2.23979	0.01789	0.00460
300	1.66330	0.05318	10.19060	1.66202	0.37313	0.00576	0.00227
900	0.60317	0.01777	4.79942	1.37207	0.24600	0.00194	0.00066
2700	0.23070	0.00594	2.50897	1.28453	0.16133	0.00066	0.00015
8100	0.10292	0.00199	1.48975	1.29131	0.06690	0.00023	0.00005

Convergence Ratios

dT	α	R^A	δ	ρ_A	σ_A	θ	κ
300/100	0.993	0.997	0.449	2.113	0.500	0.966	1.479
900/300	1.088	1.003	1.413	2.477	1.978	1.009	0.868
2700/900	1.147	1.002	1.568	2.809	1.967	1.017	0.693
8100/2700	1.338	1.004	1.781	3.016	1.244	1.063	0.927

Source: Ivashchenko and Mutschler (2020)

The `advanced=1` option of the toolbox

Point mode

Additional analysis of the LRE form (steady-state and dynamic model derivatives) for all model variables

Different Sensitivity Measure based on norm of columns of Jacobians

Analysis of identification patterns based on regressions and Singular Value Decompositions

A Different Sensitivity Measure

How do changes in the elements of θ impact the model moments, the reduced form solution, and the dynamic model

Can be measured locally using the corresponding Jacobian with certain normalizations

Norm of the columns of the standardized Jacobian yields single aggregate sensitivity measure over all moments, solution matrices, or dynamic model Jacobians for each parameter

Analyzing identification patterns

1. Check which group of one, two or more parameters is most capable to mimic (replace) the effect of each parameter

- A brute force search is done for each column of the Jacobian to detect the group of columns, having the highest explanatory power by a linear regression

2. Take singular value decomposition of Information matrix

- display eigenvectors corresponding to smallest (or highest) singular values

Example:
Advanced option in point mode

`investadjust_identif_advanced.mod`

Monte Carlo mode

Analysis of the condition number of the Jacobians of moments, reduced-form and Linear Rational Expectations model and detection of the parameters that mostly drive large condition numbers (i.e. weaker identification)

Detailed point-estimate (identification strength and collinearity analysis) of the parameters set having the smallest/largest condition number

Analysis of the identification patterns across the Monte Carlo sample

Example:
Advanced option in Monte Carlo mode

`investadjust_identif_mc_advanced.mod`

Identification of nonlinear and non-Gaussian DSGE models

Dynare's model framework

General DSGE model

$$E_t f\left(z_{t+1}, z_t, z_{t-1}, u_t \,|\, \theta\right) = 0$$

$$z_t = g(z_{t-1}, u_t \,|\, \theta)$$

Perturbation solution: Taylor-approximation around the non-stochastic steady-state:

$$z_t = \bar{z} + g_x(x_{t-1} - \bar{x}) + g_u u_t + \frac{1}{2}\left[g_{xx}(x_{t-1} - \bar{x}) \otimes (x_{t-1} - \bar{x}) + 2g_{xu}(x_{t-1} \otimes u_t) + g_{uu}(u_t \otimes u_t) + g_{\sigma\sigma}\sigma^2\right] + \frac{1}{6}[\ldots] + \ldots$$

Problem of higher-order approximations

Possibility of explosive behavior in higher-order approximations

↪ Model may not be stationary or does not have an ergodic probability distribution

Solution: Pruning

- Leave out terms in solution that have higher-order effects than the approximation order
- Kim, Kim, Schaumburg and Sims (2008) and Andreasen, Fernández-Villaverde and Rubio-Ramírez (2018) show that pruned state space is stationary and ergodic
- Lombardo and Uhlig (2017) or Lan and Meyer-Gohde (2013) provide theoretical foundation for this seemingly *ad-hoc* procedure

Univariate example

$$x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, \qquad |g_x| < 1, \qquad g_{xx} > 0$$

Two fixed-points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

↪ Once the model passes the second fixed point it explodes

Univariate example

Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx}\left(x_{t-1}^f\right)^2 + 2g_{xx}\left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx}\left(x_{t-1}^s\right)^2 + g_u u_t$$

Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$ to get law of motions:

$$x_t^f = g_x x_{t-1}^f + g_u u_t$$

$$x_t^s = g_x x_{t-1}^s + g_{xx}(x_{t-1}^f)^2$$

$$(x_t^f)^2 = g_x^2\left(x_{t-1}^f\right)^2 + 2g_x g_u x_{t-1}^f u_t + g_u^2 u_t^2$$

Univariate example

Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} x_t^f \\ x_t^s \\ x_t^{f^2} \end{pmatrix}}_{z_t} = \underbrace{\begin{pmatrix} g_x & 0 & 0 \\ 0 & g_x & g_{xx} \\ 0 & 0 & g_x^2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_{t-1}^f \\ x_{t-1}^s \\ x_{t-1}^{f^2} \end{pmatrix}}_{z_{t-1}} + \underbrace{\begin{pmatrix} g_u & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2g_x g_u & g_u^2 \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} u_t \\ x_{t-1}^f u_t \\ u_t^2 - \sigma_u^2 \end{pmatrix}}_{\xi_t} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ g_u^2 \sigma_u^2 \end{pmatrix}}_{c}$$

Pruned State Space System

Proposition:
Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned perturbation solution of a DSGE model can be rewritten as a linear time-invariant state-space system for any approximation order:

$$z_t = c + Az_{t-1} + B\xi_t$$

$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

Note: Even if u_t is Gaussian, ξ_t is not!

↪Higher-order statistics (HOS) may contain additional information for estimation and identification

Straightforward (but tedious) to compute moments, cumulants and polyspectra

Identification diagnostics for nonlinear DSGE models

Diagnostics for nonlinear DSGE

We can use similar diagnostics based on moments or spectrum to detect local identification issues

- Note that first and second moments and spectral density are computed from the pruned state-space system
- Higher-order moments and polyspectra might be also considered (not yet in Dynare)

Identification Strength

- Asymptotic information matrix is not available analytically, but covariance of moments is approximated by simulation from pruned state space
- Bayesian Learning Rate indicator is readily available; however, one needs to use a nonlinear Kalman filter or a particle filter to evaluate the likelihood

Example:
Identification via second-order approximation

`investadjust_identif_second_order.mod`

Concluding Remarks

Concluding Remarks

Fix obvious identification failures: calibrate or (better) re-parameterize your model

Use readily available tools to get insight into the workings of your model

Identifiability is a model property that depends on the choices a modeler makes

- choice of observables
- functional specifications
- model features
- choice of structural shocks

Concluding Remarks

Some personal experience

- Larger models tend to be theoretically identified, but suffer from weak identification
- Robust inference under possible weak identification (Dufour et al., 2009, 2013; Kleibergen and Mavroeidis, 2009; Mavroeidis, 2005, 2010; Guerron-Quintana et al.,2013; Andrews and Mikusheva, 2014; Qu, 2014)
- wide confidence intervals for weakly identified parameters can be accompanied by narrow bands for IRFs
- Nonlinear or non-Gaussian approach might enrich identifiability and model dynamics (but comes at a price)