Identification Analysis of DSGE models with DYNARE

Marco Ratto

European Commission
Joint Research Centre

DYNARE Summer School
Paris, June 14th 2018

Outline

1. Identification in DSGE models
   - Local identification
2. DYNARE Implementation
   - Identification strength
   - Analyzing identification patterns
   - Main features of the software
   - DYNARE procedure
   - Notes
   - Other related works
3. Examples
   - Kim (2003)
   - An Schorfheide (2007)

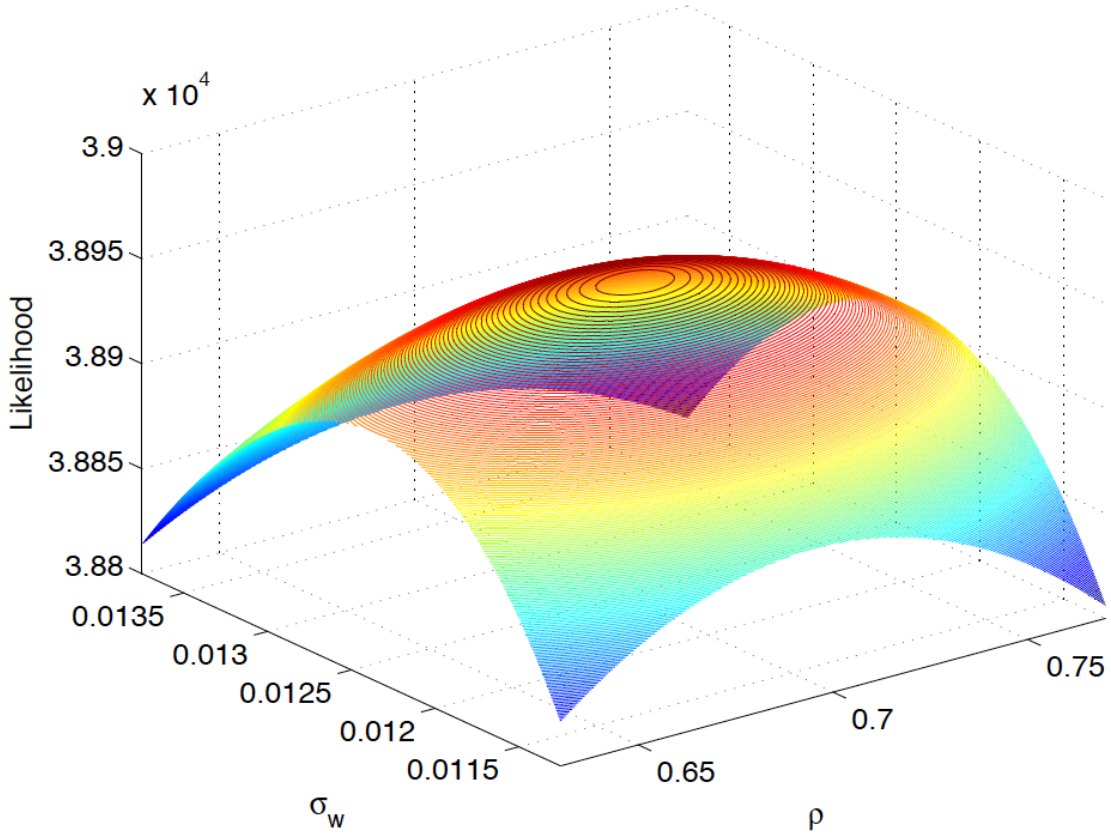
Motivation

- Yesterday you learned how to estimate parameters by evaluating the posterior $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$
- Today we formally how to verify whether all elements of $\boldsymbol{\theta}$ are properly identified

Identification in DSGE models

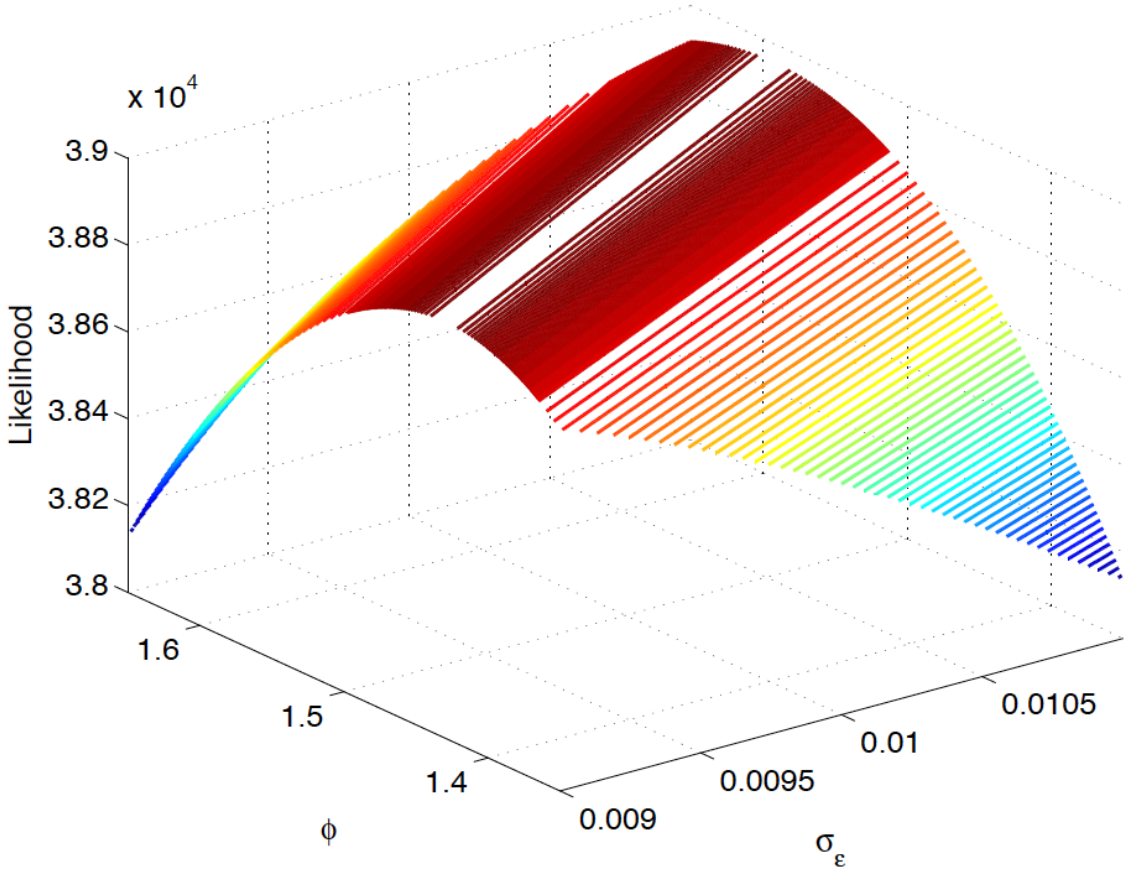
Many authors addressed identification issues in economic modeling (Canova and Sala, 2009; Iskrev, 2010b; Komunjer and Ng, 2011, 2012; Qu and Tkachenko, 2012; Mutschler, 2014).

Likelihood shape



The likelihood function shows a clear peak in ρ and σ_{wt}.

Likelihood shape flat



The likelihood function has no unique maximum in σ_ε and ψ, only a range of maxima

Identification Problems (Canova and Sala, 2009)

1. **Observational Equivalence**: mapping between structural parameters and objective function has no unique maximum
   - structural models with potentially different economic interpretations may be indistinguishable
2. **Under-identification**: objective function is independent of certain structural parameters, e.g. because they disappear from rational expectations solution
   - **Partial identification** with two structural parameters entering objective function only proportionally, making them separately unrecoverable, is a special case
3. **Weak identification**: parameter theoretically identified, but curvature may be small in certain regions of the parameter space

State Space System: Transition Equation

Is there a systematic way to detect such issues for elements of θ?

- Consider generic linearized model in state-space form where $\mathbf{z}_t = \hat{\mathbf{z}}_t - \mathbf{z}^*$

$$\mathbf{z}_t = \mathbf{A}(\theta)\mathbf{z}_{t-1} + \mathbf{B}(\theta)\mathbf{u}_t, \mathbf{u}_t \sim \mathcal{N}(0, \Sigma_{\mathbf{u}})$$

- Define $\Omega = B\Sigma_{\mathbf{u}}B'$ and the vector

$$\tau := [\tau_z', \ \tau_A', \ \tau_\Omega']'$$

as the collection of non-constant elements of $\mathbf{z}^*$, $\mathrm{vec}(\mathbf{A})$, and $\mathrm{vech}(\Omega)$, respectively, that depend on θ.

- Drops, e.g., steady states that are always zero or calibrated values
- The unconditional first and second moments are given by:

$$\begin{aligned} E(\mathbf{z}_t) &= 0 \\ E(\mathbf{z}_t\mathbf{z}_t') &\equiv \Sigma_{\mathbf{z}}(0) = \mathbf{A}\Sigma_{\mathbf{z}}(0)\mathbf{A}' + \Omega \end{aligned}$$

where the latter is the fixed point of the Lyapunov-equation

State Space System: Observation Equation

- The observables are given by

$$\mathbf{x}_t = \mathbf{C}(\mathbf{z}_t + \mathbf{z}^*)$$

- In general, only observing $\mathbf{x}_t$ would be insufficient to fully characterize the distribution of $\mathbf{z}_t$
- Fortunately, our model implies restrictions through $\boldsymbol{\theta}$.

Note: mapping from structural parameters into the above law of motion for $\mathbf{z}$ can be highly nonlinear.

The Likelihood function

- The unconditional first and second moments of the observables $\mathbf{x}_t$ are given by:

$$\mathrm{E}\,\mathbf{x}_t := \boldsymbol{\mu}_{\mathbf{x}} = \mathbf{C}\mathbf{z}^* \tag{1}$$

$$\mathrm{cov}(\mathbf{x}_{t+i}, \mathbf{x}_t') := \Sigma_{\mathbf{x}}(i) = \begin{cases} \mathbf{C}\Sigma_{\mathbf{z}}(0)\mathbf{C}' & \text{if } i = 0 \\ \mathbf{C}\mathbf{A}^i\Sigma_{\mathbf{z}}(0)\mathbf{C}' & \text{if } i > 0 \end{cases} \tag{2}$$

- stack the unique elements of $\Sigma_{\mathbf{x}}(i)$ into

$$\boldsymbol{\sigma}_T := [\mathrm{vech}(\Sigma_{\mathbf{x}}(0))', \mathrm{vec}(\Sigma_{\mathbf{x}}(1))', ..., \mathrm{vec}(\Sigma_{\mathbf{x}}(T-1))']'$$

  and define the vector of first anf second order moments

$$\mathbf{m}_T := [\boldsymbol{\mu}', \boldsymbol{\sigma}_T^{'}]'$$

- note 1: $\mathbf{m}_T$ is a function of $\boldsymbol{\theta}$.
- note 2: if either $\mathbf{u}_t$ is Gaussian, or there are no distributional assumptions about the structural shocks, identification can be expressed in terms theoretical moments up to order two only.

Definitions

Definition (Global identification)

Let $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k$ be the parameter vector of interest, and suppose that inference about $\boldsymbol{\theta}$ is made on the basis of the observed random vector $\mathbf{x}$ with a known joint probability density function $f(\mathbf{x};\boldsymbol{\theta})$. A point $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$ is said to be **globally identified** if

$$f(\mathbf{x};\tilde{\boldsymbol{\theta}}) = f(\mathbf{x};\boldsymbol{\theta}_0),\ \forall \mathbf{x} \Rightarrow \tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_0,\ \text{with probability 1} \qquad (3)$$

for any $\tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$.

Definition (Local identification)

If (3) is true only for values $\tilde{\boldsymbol{\theta}}$ in an open neighborhood of $\boldsymbol{\theta}_0$, then $\boldsymbol{\theta}_0$ is said to be **locally identified**.

The linear-Gaussian case

Theorem

Suppose our model in state-space form with $\boldsymbol{\theta}_0$ is the data generating process $\mathbf{x}_T$. Then $\boldsymbol{\theta}_0$ is globally identified if

$$\mathbf{m}_T(\tilde{\boldsymbol{\theta}}) = \mathbf{m}_T(\boldsymbol{\theta}_0) \Leftrightarrow \tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 \tag{4}$$

for any $\tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$ (and correspondingly for local identification).

- requires mapping from population moments $\mathbf{m}_T$ to $\boldsymbol{\theta}$ to be unique.
- If not satisfied, there exist parameter values that imply the same joint distribution
  $\Rightarrow$ Even with infinitely many observations, the true value could not be identified
- Checking global identification is hard $\Rightarrow$ check local identification for relevant parameter range

Local identification: The rank condition

Theorem

Suppose that $\mathbf{m}_T$ *is a continuously differentiable function of* $\boldsymbol{\theta}$*. Then* $\boldsymbol{\theta}_0$ *is locally identifiable if the Jacobian matrix* $J(q) := \frac{\partial \mathbf{m}_q}{\partial \boldsymbol{\theta}'}$ *has a full column rank at* $\boldsymbol{\theta}_0$ *for* $q \leq T$*. This condition is both necessary and sufficient when* $q = T$ *if* $\mathbf{u}_t$ *is normally distributed.*

- We need at least as many moments as parameters
- One would have to check this condition everywhere $\Rightarrow$ impossible
- Nevertheless helpful in detecting observational equivalence (column of zeros) and underidentification (linear dependence of columns)

Decompose the Jacobian

$$J(T) = \frac{\partial \mathbf{m}_T}{\partial \boldsymbol{\tau}'} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'} \tag{5}$$

- $H(T) = \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'}$ shows how the parameters $\boldsymbol{\theta}$ affect the non-constant model solution parts $\boldsymbol{\tau}$;
- $J1(T) = \frac{\partial \mathbf{m}_T}{\partial \boldsymbol{\tau}'}$ shows how the model solution maps into the observed data moments

Corollary

The point $\boldsymbol{\theta}_0$ is locally identifiable only if $H(T)$ at $\boldsymbol{\theta}_0$ has full rank.

- Necessary condition as parameters only affect distribution of observables through their effect on model solution
- It is not sufficient unless all states are observed
- $J1(T)$ provides information on whether a theoretically identifiable parameter can be identified using the observables contained in $\mathbf{x}$

The Information Matrix Revisited

- The precision of the parameter estimates at the mode was given (asymptotically) by the inverse of the **Fisher Information Matrix**

$$I(\theta) = E\left[\left(\frac{\partial \log\left(p\left(Y|\theta\right)\right)}{\partial \theta'}\right)' \left(\frac{\partial \log\left(p\left(Y|\theta\right)\right)}{\partial \theta'}\right)\right] \tag{21}$$

- Typically, non-singularity of this matrix is sufficient for local identification (for details see Iskrev 2011)
- Can be decomposed into a variance matrix Δ and a correlation matrix $\tilde{I}(\theta)$

$$I(\theta) = \Delta^{\frac{1}{2}} \tilde{I}(\theta) \Delta^{\frac{1}{2}} \tag{22}$$

where

$$\Delta = \text{diag}\left(I(\theta)\right)$$

Identification Strength

$$I(\theta) = \Delta^{\frac{1}{2}} \tilde{I}(\theta) \Delta^{\frac{1}{2}} \tag{22}$$

- Here, we can again see our two reasons for non-identifiability:
  1. The likelihood does not change when parameter i changes, i.e.

$$\Delta_i = 0 \tag{23}$$

  2. The effect on the likelihood is offset due to perfect correlation, i.e.

$$\rho_i \equiv corr\left(\frac{\partial \log(p(Y|\theta))}{\partial \theta_i}, \frac{\partial \log(p(Y|\theta))}{\partial \theta_{-i}}\right) = 1 \tag{24}$$

- This suggests that we can use this curvature information also for **weak identification** as in that case $\Delta_i \approx 0$ and/or $\rho_i \approx 1$

Identification Strength

- Identification strength can be measured using

$$s_i(\theta) = \sqrt{\Delta_i\left(1-\rho_i^2\right)} \tag{25}$$

- If $I(\theta)$ is not singular, $1/s_i(\theta)$ is equal to the square root of the ith diagonal element of $(I(\theta))^{-1}$, i.e.

$$s_i(\theta) = \frac{1}{\left((I(\theta))^{-1}\right)_{(ii)}} \tag{26}$$

- As this 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

- Problem: measure is percentage change in likelihood for unit change in parameter $\rightarrow$ not unit free
- Solution: multiply with value of θ_i to obtain elasticity measure:

$$s_i^r(\theta) = \sqrt{\theta_i \Delta_i} \times \sqrt{(1 - \rho_i^2)} \tag{27}$$

where the first part is the elasticity of the likelihood function w.r.t. θ_i, keeping all other parameters constant

Dynare Implementation

- The strength of identification for parameter θ_i is defined through a renormalized version of the curvature as

$$s_i^{dyn} = \theta_i \times s_i(\theta) \overset{(26)}{=} \sqrt{\frac{\theta_i^2}{\left(I(\theta)^{-1}\right)_{(i,i)}}}$$

- Can be interpreted as an "a-priori t-test"
- Taking square of θ assures it to be positive
- The **sensitivity component** contained in this measure (as opposed to the **correlation component**) is defined as

$$\tilde{\Delta}_i = \sqrt{\theta_i^2 \Delta_i} = \sqrt{\theta_i^2 I(\theta)_{(i,i)}}$$

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

Dynare Implementation

- Because some parameters may have a prior mean of zero, an alternative normalization uses the prior standard deviation $\sigma(\theta_i)$:

$$s_i = \sqrt{\frac{\left(\sigma\left(\theta_i\right)\right)^2}{\left(I(\theta)^{-1}\right)_{(i,i)}}}$$

$$\tilde{\Delta}_i = \sqrt{\left(\sigma\left(\theta_i\right)\right)^2 I(\theta)_{(i,i)}}$$

DYNARE implementation

The identifiability of each draw θ_j is then established using the necessary and sufficient conditions discussed by Iskrev (2010b):

- Finding that matrix H is rank deficient at θ_j implies that this particular point in Θ is unidentifiable in the model.
- Finding that H has full rank but $J(T)$ does not, means that θ_j cannot be identified given the set of observed variables and the number of observations.
- if θ is identified check $J(q)$ with $q < T$. According to Theorem 2 this is sufficient for identification;
- with these tests it is possible to detect identification problems that are inherent to the structure of the DSGE model, *before* taking the model to the data!

Tracking singularities

Whenever some of the matrices H, $J(T)$ is rank deficient, the code tries to diagnose the subset of parameters responsible for the rank deficiency:

1. if there are columns of zeros in the $J(\cdot)$ matrix, the associated parameter is printed on the MATLAB command window;
2. compute pairwise- and multi-correlation coefficients for each column of the $J(\cdot)$ matrix: if there are parameters with correlation coefficients equal to unity, these are printed on the MATLAB command window;
3. take the Singular Values Decomposition (SVD) of $J(\cdot)$ and track the eigenvectors associated to the zero singular values.

Asymptotic Information Matrix. Given a sample size T, the Fischer information matrix $\mathcal{I}_T(\boldsymbol{\theta})$ is computed as discussed in Iskrev (2010a). This also provides an estimate of the Cramer-Rao lower bound of uncertainty of $\boldsymbol{\theta}$.

Simulated moments. perform stochastic simulations for T periods and compute sample moments of observed variables; repeat for N_r replicas, getting N_r samples of simulated moments; take the covariance matrix $\Sigma(\mathbf{m}_T)$ of (first and second) simulated moments. A 'moment information matrix' can be defined as $\mathcal{I}_T(\boldsymbol{\theta}|\mathbf{m}_T) = J(T) \cdot \Sigma(\mathbf{m}_T)^{-1} \cdot J(T)'$. This does NOT provide an estimate of the Cramer-Rao lower bound. Used with stochastic singularity and whenever the Asymptotic Information Matrix cannot be computed.

Analyzing identification patterns

1. Andrle (2010): the identification patterns are shown by taking the singular value decomposition of $\mathcal{I}_T(\boldsymbol{\theta})$ or of the $J(q)$ matrix and displaying the eigenvectors corresponding to the smallest (or highest) singular values;
2. Iskrev (2010b): 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 $J(q)_{(j)}$ to detect the group of columns $J(q)_{(I \not\ni j)}$, having the highest explanatory power for $J(q)_{(j)}$ by a linear regression.

Main features of the software

The new DYNARE keyword `identification` triggers the routines developed at JRC. This option has two modes of operation; point identification check (default) and Monte Carlo mode.

Point identification check

With a list of estimated parameters:

- with prior definition: the program performs the local identification checks for the estimated parameters at the prior mean (prior mode, posterior mean and posterior mode are also alternative options);
- for ML estimation (no prior definition), local identification checks are performed for the estimated parameters at the actual or initial value declared for estimation (ML value is also possible);

No list of estimated parameters:

- the program computes the local identification checks for *all* the model parameter values declared in the DYNARE model file.

Monte Carlo exploration

- draw θ from Θ (ensuring stability and determinacy), compute rank and condition number of H and $J(T)$, and repeat this many times
- if H is rank-deficient at θ_j, this particular point is unidentifiable
- if H has full rank but $J(T)$ does not, then θ_j cannot be identified for the particular set of observables and contemporaneous and lagged moments under consideration, i.e. given vx^T and T

The Advanced option (point)

- analysis of the LRE form and the reduced form;
- identification patterns (Iskrev, 2010b; Andrle, 2010).

The Advanced option (Monte Carlo)

- analysis of the condition number of the Jacobians $J(q)$, H and J_Γ and detection of the parameters that mostly drive large condition numbers (i.e. weaker identification);
- analysis of the identification patters across the Monte Carlo sample;
- detailed point-estimate (identification strength and collinearity analysis) of the parameters set having the smallest/largest condition number;
- when some singularity (rank condition failure) is detected for some elements of the Monte Carlo sample, detailed point-estimates are performed for such critical points.

Test routines

A library of test routines is also provided in the GiT repository. Such tests implement:

Kim (2003) : the DYNARE routines for this example are placed in the folder `dynare_root/tests/identification/kim`;

An and Schorfheide (2007) : the DYNARE routines for this example are placed in `dynare_root/tests/identification/as2007`;

DYNARE procedure

DYNARE keyword

`identification(<options>=<values>);`

triggers the point local identification checks, performed at the prior mean. Prior definitions and the list of observed values are needed, using the standard DYNARE syntax for setting-up an estimation.

DYNARE procedure: Options I

- `parameter_set = prior_mode | prior_mean | posterior_mode | posterior_mean | posterior_median`. Specify the parameter set to use. Default: `prior_mean`.
- `prior_mc = INTEGER` sets the number of Monte Carlo draws (default = 1); `prior_mc=1` triggers the default point identification analysis; `prior_mc>1` triggers the Monte Carlo mode;
- `prior_range = INTEGER` triggers uniform sample within the range implied by the prior specifications (when `prior_mc>1`). Default: 0
- `load_ident_files = 0`, triggers a new analysis, while `load_ident_files = 1`, loads and displays a previously performed analysis (default = 0);

DYNARE procedure: Options II

- `ar = <integer>` (default $= 3$), triggers the value for q in computing $J(q)$;
- `useautocorr`: this option triggers $J(q)$ in the form of auto-covariances and cross-covariances (`useautocorr = 0`), or in the form of auto-correlations and cross-correlations (`useautocorr = 1`). The latter form normalizes all $\mathbf{m}_q$ entries in $[-1, 1]$ (default $= 0$).
- `advanced = INTEGER` triggers standard or advanced identification analysis (default $= 0$).
- `max_dim_cova_group = INTEGER` In the brute force search (performed when `advanced=1`) this option sets the maximum dimension of groups of parameters that best reproduce the behavior of each single model parameter. Default: 2

DYNARE procedure: Options III

- `periods = INTEGER` triggers the length of the stochastic simulation to compute the analytic Hessian. Default: 300
- `periods = INTEGER` When the analytic Hessian is not available (i.e. with stochastic singularity or diffuse Kalman filter), this triggers the length of stochastic simulation to compute Simulated Moments Uncertainty. Default: 300
- `replic = INTEGER` When the analytic Hessian is not available, this triggers the number of replicas to compute Simulated Moments Uncertainty. Default: 100.
- `gsa_sample_file = INTEGER` If equal to 0, do not use sample file. If equal to 1, triggers GSA prior sample. If equal to 2, triggers GSA Monte-Carlo sample (i.e. loads a sample corresponding to `pprior`=0 and `ppost`=0 in the `dynare_sensitivity` options). Default: 0

DYNARE procedure: Options IV

- `gsa_sample_file = FILENAME` Uses the provided path to a specific user defined sample file. Default: 0

Notes

- DYNARE 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 in the `<>_steadystate.m` file.
- The # syntax should be used in the `model` block of the DYNARE file, instead.

Numerical Issues

- Detection of perfect collinearity among derivatives of theoretical moments of the observables with respect to different deep parameters:
- key issue: distinguish possible weak indentification, that is, near linear dependence, form true perfect collinearity.
- 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.
- clear advantage in the analytic derivation.

Other related works

Komunjer and Ng (2011):

- use structural properties of the canonical solution of DSGE models and restrictions implied by observational equivalence to derive rank conditions that do not require knowledge of infinite autocovariances or Markov parameters;
- like Iskrev, their conditions do not depend on data but also do not depend on the estimator (i.e. do not rely on Gaussian assumptions about shocks) nor on T (i.e. apply to $T \to \infty$);
- useful framing in terms of controllability/observability and minimal state space representation (control system engineering);
- no big practical advance w.r.t. the present approach for identification purposes, but useful additional diagnostics to analyse DSGE models.
- use numerical derivatives.

Spectral Density

- Komunjer and Ng (2011) showed that the DSGE model can be written as an MA(∞)

$$x_t = \left(C(\theta) \left(I - A(\theta) L \right)^{-1} B(\theta) L \right) u_t \tag{8}$$

  where L is the lag operator.
- Two vectors in the parameter space are observationally equivalent if the two system have the same spectral density

Necessary and sufficient condition for local identification

- Two triples $(\theta_0, I_{n_x}, I_{n_e})$ and (θ_1, T, U) are observationally equivalent if

$$A(\theta_1) = TA(\theta_0)T^{-1} \quad B(\theta_1) = TB(\theta_0)U \quad C_j(\theta_1) = C_j(\theta_0)T^{-1}$$
$$\Sigma(\theta_1) = U^{-1}\Sigma(\theta_0)U^{-1}$$

  with T and U being full rank matrices.

- A necessary and sufficient condition for local identification at θ_0 is that

$$rank(\Delta_j(\theta_0)) = n_\theta + n_x^2 + n_e^2 \tag{9}$$

  where

$$\Delta_j(\theta_0) \equiv \Delta_j(\theta_0, I_{n_x}, I_{n_e}) = \left(\frac{\partial \delta_j(\theta_0, I_{n_x}, I_{n_e})}{\partial \theta}, \frac{\partial \delta_j(\theta_0, I_{n_x}, I_{n_e})}{\partial T}, \frac{\partial \delta_j(\theta_0, I_{n_x}, I_{n_e})}{\partial U} \right)$$
$$\equiv \left(\Delta_{j,\Lambda}(\theta_0),\ \Delta_{j,T}(\theta_0),\ \Delta_{j,U}(\theta_0) \right)$$

Other related works

Weak Identification:Guerron-Quintana et al. (2013), Doufour et al (2013), Andrews Mikusheva (2015), Qu (2015)

- Parameters are point-identified but classical inference (bands and testing) is blurred.
- weak parameter identification results in the failure of classical asymptotic to provide good approximation to finite sample distributions of test statistics.

Prior Informativeness and Sensitivity: Muller (2012).

Examples

Kim (2003)

This paper demonstrated a functional equivalence between two types of adjustment cost specifications, coexisting in macroeconomic models with investment: intertemporal adjustment costs which involve some degree of complementarity between capital and investment in capital accumulation, and multi-sectoral costs which are captured by a nonlinear transformation between consumption and investment.

We reproduce results of Kim (2003), worked out analytically, applying the DYNARE procedure on the non-linear form of the model.

The representative agent maximizes

$$\sum_{t=0}^{\infty} \beta^t \log C_t \tag{10}$$

subject to a national income identity and a capital accumulation equation:

$$(1-s)\left(\frac{C_t}{1-s}\right)^{1+\theta} + s\left(\frac{I_t}{s}\right)^{1+\theta} = (A_t K_t^{\alpha})^{1+\theta} \tag{11}$$

$$K_{t+1} = \left[\delta\left(\frac{I_t}{\delta}\right)^{1-\phi} + (1-\delta)K_t^{1-\phi}\right]^{\frac{1}{1-\phi}} \tag{12}$$

where $s = \frac{\beta\delta\alpha}{\Delta}$, $\Delta = 1 - \beta + \beta\delta$. when $\phi \to 0$ and $\theta \to 0$, we have the usual linear capital accumulation technology and linear resource constraint.

- $\phi(\geq 0)$ is the inverse of the elasticity of substitution between I_t and K_t (represents the size of intertemporal adjustment costs)
- $\theta(\geq 0)$ is the inverse of the elasticity of transformation between consumption and investment (called the multisectoral adjustment cost parameter).
- the two adjustment cost parameter only enter through an 'overall' adjustment cost parameter $\Phi = \frac{\phi+\theta}{1+\theta}$, thus implying that they cannot be identified separately.

On the effect of the number of states.

Assume to have an additional equation for the Lagrange multiplier $\lambda_t = \frac{(1-s)^\theta}{(1+\theta)C_t^{(1+\theta)}}$, with λ_t entering the Euler equation.

We still assume that only C_t and I_t can be observed.

An and Schorfheide (2007)

$$y_t = \mathbb{E}_t[y_{t+1}] + g_t - \mathbb{E}_t[g_{t+1}] - 1/\tau \cdot (R_t - \mathbb{E}_t[\pi_{t+1}] - \mathbb{E}_t[z_{t+1}]) \quad (13)$$

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa(y_t - g_t) \quad (14)$$

$$(R_t = \rho_R R_{t-1} + (1-\rho_R)\psi_1 \pi_t + (1-\rho_R)\psi_2(\Delta y_t + z_t) + \varepsilon_{R,t})$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\psi_1 \pi_t + (1-\rho_R)\psi_2(y_t - g_t) + \varepsilon_{R,t} \quad (15)$$

$$g_t = \rho_g g_{t-1} + \varepsilon_{g,t} \quad (16)$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t} \quad (17)$$

where y_t is GDP in efficiency units, π_t is inflation rate, R_t is interest rate, g_t is government consumption and z_t is change in technology.

Taylor rule is not identified using $y_t - g_t$, while it is using $(\Delta y_t + z_t)$.

The model is completed with three observation equations for quarterly GDP growth rate (YGR_t), annualized quarterly inflation rates (INF_t) and annualized nominal interest rates (INT_t):

$$YGR_t = \gamma^Q + 100 * (y_t - y_{t-1} + z_t) \tag{18}$$

$$INFL_t = \pi^A + 400\pi_t \tag{19}$$

$$INT_t = \pi^A + r^A + 4\gamma^Q + 400R_t \tag{20}$$

where $\beta = \frac{1}{1+r^A/400}$.

Komunjer and Ng (2011) study its identification failure:

- they associate the lack of identification with linear dependence among the three Taylor rule parameters, but exact collinearity occurs only when the variance of the monetary shock is included
- they set much larger level of tolerance for Matlab's rank in order to correctly detect lack of identification (1.e-3).

We study its identification failure with DYNARE:

- Using analytic derivatives, DYNARE provides the right answer with TOL=1.e-13 (and larger) which is also the default in DYNARE, while using normalization (which is the default in DYNARE Identification) TOL can be tightened at 1.e-17.
- This example shows that, when non-identification is due to collinearity, numerical derivatives may imply a sensitivity with respect to numerical tolerances in determining the rank in Jacobians.
- This can have strong implications in the understanding of the sources of non-identification.

Bibliography I

Sungbae An and Frank Schorfheide. Bayesian analysis of DSGE models. *Econometric Reviews*, 26(2-4):113–172, 2007. DOI:10.1080/07474930701220071.

Michal Andrle. A note on identification patterns in DSGE models (august 11, 2010). ECB Working Paper 1235, 2010. Available at SSRN: http://ssrn.com/abstract=1656963.

Fabio Canova and Luca Sala. Back to square one: identification issues in DSGE models. *Journal of Monetary Economics*, 56(4), May 2009.

Nikolay Iskrev. Evaluating the strenght of identification in DSGE models. an a priori approach. unpublished manuscript, 2010a.

Nikolay Iskrev. Local identification in DSGE models. *Journal of Monetary Economics*, 57:189–202, 2010b.

Bibliography II

Jinill Kim. Functional equivalence between intertemporal and multisectoral investment adjustment costs. *Journal of Economic Dynamics and Control*, 27(4):533–549, February 2003. URL http://ideas.repec.org/a/eee/dyncon/v27y2003i4p533-549.html.

Ivana Komunjer and Serena Ng. Dynamic identification of DSGE models. *Econometrica*, 79(6):1995–2032–1423, November 2011.

Ivana Komunjer and Serena Ng. Global identification in nonlinear models with moment restrictions. *Econometric Theory*, 28: 719–729, 2012.

Bibliography III

Willi Mutschler. Identification of DSGE Models - the Effect of Higher-Order Approximation and Pruning. CQE Working Papers 3314, Center for Quantitative Economics (CQE), University of Muenster, October 2014. URL http://ideas.repec.org/p/cqe/wpaper/3314.html.

Zhongjun Qu and Denis Tkachenko. Identification and frequency domain quasimaximum likelihood estimation of linearized dynamic stochastic general equilibrium models. *Quantitative Economics*, 3(1):95–132, 03 2012. URL http://ideas.repec.org/a/ecm/quante/v3y2012i1p95-132.html.

Computing the Jacobian

- Numerical differentiation: tends to be inaccurate for highly non-linear functions, possibly leading to wrong conclusions concerning the rank of the Jacobian matrix and the identifiability of the parameters in the model.

- Iskrev (2010b) applied analytical derivatives, employing implicit derivation, exploiting

$$J(T) = \frac{\partial \mathbf{m}_T}{\partial \boldsymbol{\tau}'} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'} \tag{21}$$

- The derivation of $J_1(T) := \frac{\partial \mathbf{m}_T}{\partial \boldsymbol{\tau}'}$ is straightforward, given the derivatives of $\Sigma_z(0)$ from the Lyapunov equation:

$$d\Sigma_z(0) = \mathbf{A} d\Sigma_z(0) \mathbf{A}' + \left(d\boldsymbol{\Omega} + d\mathbf{A}\Sigma_z(0)\mathbf{A}' + \mathbf{A}\Sigma_z(0)d\mathbf{A}' \right) \tag{22}$$

- The elements of the second term $J_2(T) := \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'}$, the Jacobian of the transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\tau}$, can be divided into three groups corresponding to the three blocks of $\boldsymbol{\tau}$: $\boldsymbol{\tau}_z$, $\boldsymbol{\tau}_A$ and $\boldsymbol{\tau}_\Omega$.

Computing the Jacobian, τ_z

$\hat{\mathbf{z}}^*$ is in general not known as an explicit function of $\boldsymbol{\theta}$.

The 'static' model $\mathbf{g}^* = \mathbf{g}(\hat{\mathbf{z}}^*, \hat{\mathbf{z}}^*, \hat{\mathbf{z}}^*, 0|\boldsymbol{\theta}) = 0$ provides and implicit function between $\hat{\mathbf{z}}^*$ and $\boldsymbol{\theta}$:

$$\frac{\partial \hat{\mathbf{z}}^*}{\partial \boldsymbol{\theta}'} = -\left(\frac{\partial \mathbf{g}^*}{\partial \hat{\mathbf{z}}^{*\prime}}\right)^{-1} \cdot \frac{\partial \mathbf{g}^*}{\partial \boldsymbol{\theta}'} \tag{23}$$

$\frac{\partial \boldsymbol{\tau}_z}{\partial \boldsymbol{\theta}'}$ is obtained by removing the zeros corresponding to the constant elements of $\hat{\mathbf{z}}^*$.

Computing the Jacobian, τ_A τ_Ω

Considering the general structural form:

$$\Gamma_0(\theta, \hat{\mathbf{z}}^*)\mathbf{z}_t = \Gamma_1(\theta, \hat{\mathbf{z}}^*)\,\mathrm{E}_t\,\mathbf{z}_{t+1} + \Gamma_2(\theta, \hat{\mathbf{z}}^*)\mathbf{z}_{t-1} + \Gamma_3(\theta, \hat{\mathbf{z}}^*)\mathbf{u}_t$$

The derivatives of Γ_i are:

$$\frac{\partial \Gamma_i(\theta)}{\partial \theta'} = \frac{\partial \Gamma_i(\theta, \hat{\mathbf{z}}^*(\theta))}{\partial \theta'} = \frac{\partial \Gamma_i(\theta, \hat{\mathbf{z}}^*)}{\partial \theta'} + \frac{\partial \Gamma_i(\theta, \hat{\mathbf{z}}^*)}{\partial \hat{\mathbf{z}}^{*\prime}} \cdot \frac{\partial \hat{\mathbf{z}}^*}{\partial \theta'} \quad (24)$$

Computing the Jacobian, τ_A τ_Ω

An implicit function of $\boldsymbol{\theta}$ and $\mathrm{vec}(\mathbf{A})$ is provided by the restrictions the structural model imposes on the reduced form:

$$(\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\mathbf{A})\mathbf{z}_t = \boldsymbol{\Gamma}_2\mathbf{z}_{t-1} + \boldsymbol{\Gamma}_3\mathbf{u}_t \tag{25}$$

and

$$\Big(\boldsymbol{\Gamma}_0(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\mathbf{A}\Big)\mathbf{A} - \boldsymbol{\Gamma}_2(\boldsymbol{\theta}) = \mathbf{O} \tag{26}$$

Vectorizing (26) and applying the implicit function theorem gives

$$\frac{\partial \mathrm{vec}(\mathbf{A})}{\partial \boldsymbol{\theta}'} = -\left(\frac{\partial \mathrm{vec}(\mathbf{F})}{\partial \mathrm{vec}(\mathbf{A})'}\right)^{-1} \frac{\partial \mathrm{vec}(\mathbf{F})}{\partial \boldsymbol{\theta}'} \tag{27}$$

Computing the Jacobian, τ_A τ_Ω

Closed-form expressions for computing the derivatives in (27) are provided in Iskrev (2010b). Such a derivation requires the use of Kronecker products, implying a dramatic growth in memory allocation requirements and in computational time as the size of the model increases. The typical size of matrices to be handled in Iskrev (2010b) is of $m^2 \times m^2$, which grows very rapidly with m.

Here we propose an alternative method to compute derivatives, allowing to reduce both memory requirements and the computational time.

Computing the Jacobian, $\boldsymbol{\tau}_A \quad \boldsymbol{\tau}_\Omega$

Take the derivative of (26) with respect to each θ_j, for $j = 1, \ldots, k$:

$$\mathbf{M}(\boldsymbol{\theta})\frac{\partial \mathbf{A}}{\partial \theta_j} + \mathbf{N}(\boldsymbol{\theta})\frac{\partial \mathbf{A}}{\partial \theta_j}\mathbf{P}(\boldsymbol{\theta}) = \mathbf{Q}_j(\boldsymbol{\theta}) \tag{28}$$

where

$$\begin{aligned}
\mathbf{M}(\boldsymbol{\theta}) &= \Big(\boldsymbol{\Gamma}_0(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\Big) \\
\mathbf{N}(\boldsymbol{\theta}) &= -\boldsymbol{\Gamma}_1(\boldsymbol{\theta}) \\
\mathbf{P}(\boldsymbol{\theta}) &= \mathbf{A}(\boldsymbol{\theta}) \\
\mathbf{Q}_j(\boldsymbol{\theta}) &= \frac{\partial \boldsymbol{\Gamma}_2}{\partial \theta_j} - \Big(\frac{\partial \boldsymbol{\Gamma}_0}{\partial \theta_j} - \frac{\partial \boldsymbol{\Gamma}_1}{\partial \theta_j}\mathbf{A}(\boldsymbol{\theta})\Big)\mathbf{A}(\boldsymbol{\theta})
\end{aligned}$$

Computing the Jacobian, $\boldsymbol{\tau}_A \quad \boldsymbol{\tau}_\Omega$

Equation (28) is a generalized Sylvester equation and can be solved using available algebraic solvers.

In practice we replace here the single big algebraic problem of dimension $m^2 \times m^2$ of Iskrev (2010b) with a set of k problems of dimension $m \times m$.

Computing the Jacobian, τ_A τ_Ω

From equations (25) and reduced form we have

$$\Big(\Gamma_0 - \Gamma_1\mathbf{A}\Big)\mathbf{B} = \Gamma_3 \tag{29}$$

and therefore

$$d\mathbf{B} = \Big(\Gamma_0 - \Gamma_1\mathbf{A}\Big)^{-1}\Big(d\Gamma_3 - (d\Gamma_0 - d\Gamma_1\mathbf{A} - \Gamma_1 d\mathbf{A})\Big) \tag{30}$$

Using $\Omega = \mathbf{B}\mathbf{B}'$, the differential of Ω is given by

$$d\Omega = d\mathbf{B}\mathbf{B}' + \mathbf{B}d\mathbf{B}' \tag{31}$$

Thus, once $\frac{\partial \mathrm{vec}(\mathbf{A})}{\partial \theta'}$ is available, it is straightforward to compute, first $\frac{\partial \mathrm{vec}(\mathbf{B})}{\partial \theta'}$ and $\frac{\partial \mathrm{vech}(\Omega)}{\partial \theta'}$, and then $\frac{\partial \tau_A}{\partial \theta'}$ and $\frac{\partial \tau_\Omega}{\partial \theta'}$.