A Cluster-Grid Projection Method for Solving Problems with High Dimensionality

Kenneth Judd, Lilia Maliar and Serguei Maliar

July 22, 2011

Cluster-grid algorithm (CGA)

- **A novel accurate method for solving dynamic economic models:** works for problems with high dimensionality, intractable for earlier solution methods:
  - *we solve models with hundreds of state variables using a laptop.*
- **Related literature focuses on much lower dimensionality:** a special JEDC 2011's issue compares solution methods (including our CGA) using models with 20 state variables at most.
- **Examples of potential CGA applications:**
  - macroeconomics (many heterogeneous agents);
  - international economics (many countries);
  - industrial organization (many firms);
  - finance (many assets);
  - climate change (many sectors and countries); etc.
- **CGA is a global method:** can handle strong non-linearities and inequality constraints.
  - *we solve a new Keynesian model with the zero lower bound.*

Ingredients of CGA

- **Endogenous solution domain:** our grid is constructed by clustering methods to surround the ergodic set - we avoid costs of finding a solution in the areas of state space that are never visited in equilibrium.
- **Low-cost integration:** non-product monomial and one-point quadrature integration rules.
- **Efficient solver for finding the polynomial coefficients:** fixed-point iteration.
- **Vectorized approaches for finding the control variables:** precomputation and iteration-on-allocation by Maliar, Maliar and Judd (2011).

$$\Downarrow$$

- **Taken together, these ingredients allow us to meet challenges of high-dimensional problems.**

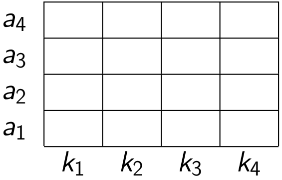
Standard projection methods, Judd (1992)

Characteristic features

- Solve a model on a prespecified grid of points.
- Use quadrature integration for approximating conditional expectations.
- Compute polynomial coefficients of policy functions using Newton's type solver.

Projection methods: curse of dimensionality

- Very accurate and fast with few state variables but cost grows exponentially with dimensionality!
  (a) Product hypercube domain $\Longrightarrow$ Curse of dimensionality!
  (b) Product quadrature integration $\Longrightarrow$ Curse of dimensionality!
  (c) Newton's solver (Jacobian, Hessian) $\Longrightarrow$ Curse of dimensionality!

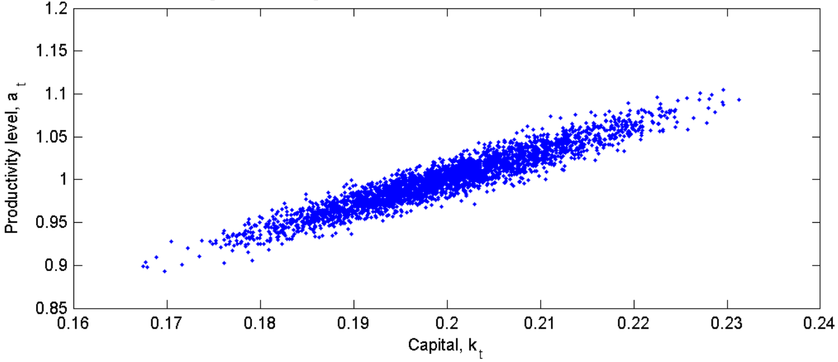


- 2 state variables with 4 grid points $\Rightarrow 4 \times 4 = 4^2 = 16$
- 3 state variables with 4 grid points $\Rightarrow 4^3 = 64$
- 10 state variables with 4 grid points $\Rightarrow 4^{10} = 1,048,576$
(With 100 grid points $\Rightarrow 100^{10} = 10^{20}$).

- *Kruger and Kubler (2004):* Smolyak's sparse grid - reduces the number of points within the multidimensional hypercube domain but not the size of the hypercube domain itself.

Ergodic set – area of the state space that is visited in simulation.
Example: time-series solution to the standard stochastic growth model with two state variables, capital and productivity

Figure 1. The ergodic set in the model with a closed-form solution.



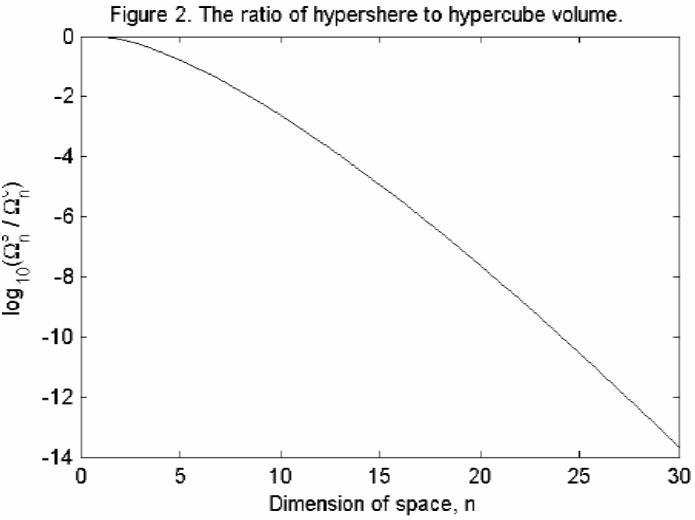
Ratio of hypersphere volume to hypercube volume

- **2-dimensional case:** a circle inscribed within a square occupies about 79% of the area of the square.
- ***n*-dimensional case:** the ratio of a hypersphere's volume Ω_n^s to a hypercube's volume Ω_n^c

$$\frac{\Omega_n^s}{\Omega_n^c} = \begin{cases} \frac{(\pi/2)^{\frac{n-1}{2}}}{1\cdot 3\cdot \ldots \cdot n} & \text{for } n = 1, 3, 5\ldots \\ \frac{(\pi/2)^{\frac{n}{2}}}{2\cdot 4\cdot \ldots \cdot n} & \text{for } n = 2, 4, 6\ldots \end{cases}.$$

- **Ratio $\frac{\Omega_n^s}{\Omega_n^c}$ declines rapidly with the dimension of the state space:**
  - when $n = 10$, the ratio $\frac{\Omega_n^s}{\Omega_n^c} = 3 \cdot 10^{-3}$;
  - when $n = 30$, the ratio $\frac{\Omega_n^s}{\Omega_n^c} = 2 \cdot 10^{-14}$.

Ergodic set versus tensor-product grid: estimated reduction in cost



Projection method on the ergodic set

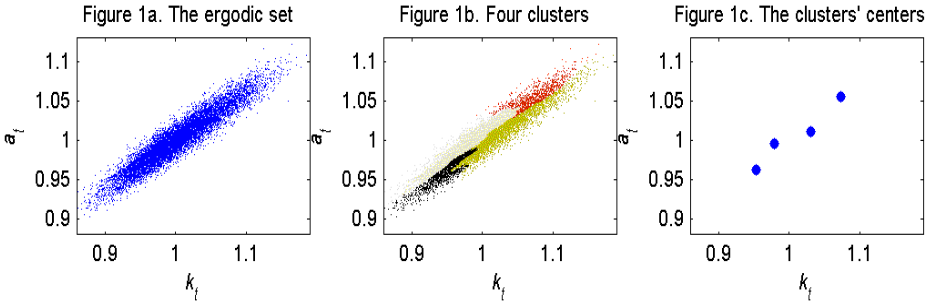
- The hypersphere ergodic set is just a tiny fraction of the hypercube tensor-product grid.
- We will develop a projection method operating on the ergodic set.
- We will construct a grid of points surrounding the ergodic set using clustering methods.

A grid of points surrounding the ergodic set

A grid of clusters' centers

1. Simulate time series solution to the model (the ergodic set), $\{k_t, a_t\}_{t=1}^{T}$.
2. Construct M clusters using methods from clustering analysis, e.g., hierarchical agglomerative or K-means clustering algorithms.
3. Compute the centers of the constructed clusters.
4. Use the clusters' centers as a grid points in multi-dimensional space.

The ergodic set and 4 clusters



Steps of the agglomerative hierarchical clustering algorithm

The zero-order partition $\mathcal{P}_0$ is the set of singletons – each observation represents a cluster.

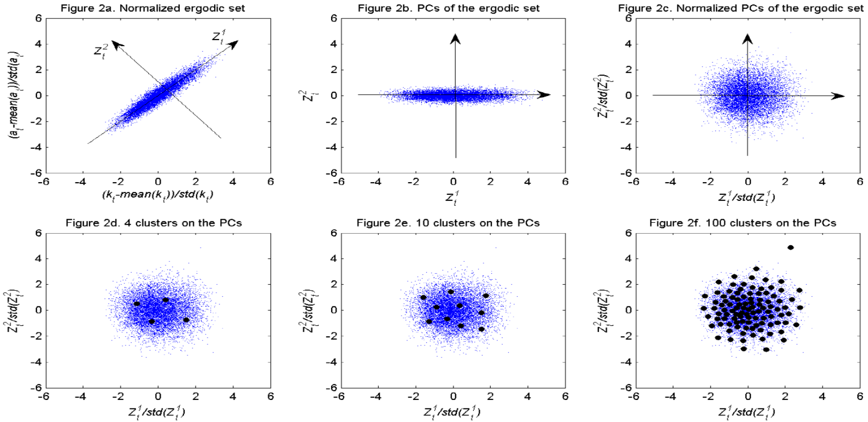
Initialization. Choose clustering linkage (we use Ward's linkage). Choose the number of clusters to be created M.

Step 1. On iteration i, compute all pairwise distances between the clusters in the partition $\mathcal{P}_i$.

Step 2. Merge a pair of clusters with the smallest distance into a new cluster. The resulting partition is $\mathcal{P}_{i+1}$.

Iterate on Steps 1 and 2. Stop when the number of clusters in the partition is M.

Clusters on principal components of the ergodic set



Properties of the cluster grid

- **The model is solved on the ergodic set** (as is done under stochastic simulation).
- **The cluster grid is more efficient than stochastic simulation:** a large number of closely-situated simulated points is replaced with a smaller number of "representative" points.
- **The cluster grid is (mostly) fixed**, while stochastic simulation algorithms redraw the simulated points on each iteration (numerical stability).
- **The cluster grid is cheap:** constructing 300 clusters on simulated series of 10,000 observations takes:
  - 9 seconds with 2 state variables
  - just 66 seconds with 200 state variables!
- **However, the cluster-grid alone does not prevent the course of dimensionality.**

Example: a representative-agent model

The representative-agent neoclassical growth model:

$$\max_{\{k_{t+1},c_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$\text{s.t.} \quad c_t + k_{t+1} = (1-\delta) k_t + a_t f(k_t),$$

$$\ln a_{t+1} = \rho \ln a_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{N}(0, \sigma^2)$$

where initial condition (k_0, a_0) is given;
$u(\cdot)$ = utility function; $f(\cdot)$ = production function;
c_t = consumption; k_{t+1} = capital; a_t = productivity;
β = discount factor; δ = depreciation rate of capital;
ρ = autocorrelation coefficient of the productivity level;
σ = standard deviation of the productivity shock.

Description of CGA

Parameterize the RHS of the Euler equation by a polynomial $\Psi(k_m, a_m; b)$,

$$\begin{aligned} k'_m &= E\left\{\beta \frac{u'(c'_m)}{u'(c_m)}\left[1-\delta+a'_m f'(k'_m)\right]k'_m\right\} \\ &\approx \Psi(k_m, a_m; b) = b_0 + b_1 k_m + b_2 a_m + ... \end{aligned}$$

Step 1. Simulate time series $\{k_t, a_t\}_{t=1}^{T+1}$ and construct M clusters. Use clusters' centers $\{k_m, a_m\}_{m=1}^{M}$ as a grid.

Step 2. Fix $b \equiv (b_0, b_1, b_2, ...)$. Given $\{k_m, a_m\}_{m=1}^{M}$ solve for $\{c_m\}_{m=1}^{M}$.

Step 3. Compute the expectation using numerical integration (quadrature integration or monomial rules)

$$\widehat{k}'_m \equiv E\left\{\beta \frac{u'(c'_m)}{u'(c_m)}\left[1-\delta+a'_m f'(k'_m)\right]k'_m\right\}.$$

Regress $\widehat{k}'_m$ on $(1, k_m, a_m, k_m^2, a_m^2, ...) \Longrightarrow$ get $\widehat{b}$.

Step 4. Solve for the coefficients using fixed-point iteration with damping,

$$b^{(j+1)} = (1-\xi)\, b^{(j)} + \xi \widehat{b}, \quad \xi \in (0,1).$$

Representative-agent model: parameter choice

- Production function: $f(k_t) = k_t^\alpha$ with $\alpha = 0.36$.
- Utility function: $u(c_t) = \frac{c_t^{1-\gamma}-1}{1-\gamma}$ with $\gamma \in \{0.2, 1, 5\}$.
- Process for shocks: $\ln a_{t+1} = \rho \ln a_t + \epsilon_{t+1}$, with $\rho \in \{0.95, 0.99\}$ and $\sigma \in \{0.01, 0.03\}$.
- Discount factor: $\beta = 0.99$.
- Depreciation rate: $\delta = 0.025$.
- Accuracy is measured by an Euler-equation error,

$$\mathcal{E}(k_t, a_t) \equiv E_t\left[\beta \frac{c_{t+1}^{-\gamma}}{c_t^{-\gamma}}\left(1 - \delta + \alpha a_{t+1} k_{t+1}^{\alpha-1}\right)\right] - 1$$

Table 1. Accuracy and speed in the representative-agent model

Polynomial degree	Mean error	Max error	CPU (sec)
1st degree	−4.32	−3.68	11.59
2nd degree	−6.12	−5.46	0.30
3rd degree	−7.58	−6.93	0.26
4th degree	−8.91	−7.87	0.14
5th degree	−9.99	−8.85	0.24

Mean and Max are unit-free Euler equation errors in log10 units, e.g.,

- -4 means $10^{-4} = 0.0001 \;\; (0.01\%)$;
- -4.5 means $10^{-4.5} = 0.0000316 \;\; (0.00316\%)$.

Benchmark parameters: $\delta = 0.025$, $\gamma = 1$, $\rho = 0.95$, $\sigma = 0.01$.

In the paper, many parameterizations are explored:

- low risk aversion: $\gamma = 1/5$;
- high risk aversion: $\gamma = 5$;
- highly persistent shocks: $\rho = 0.99$;
- highly volatile shocks: $\sigma = 0.03$.

Multi-country model

The planner maximizes a weighted sum of N countries' utility functions:

$$\max_{\left\{\left\{c_t^h, k_{t+1}^h\right\}_{h=1}^N\right\}_{t=0}^\infty} E_0 \sum_{h=1}^N v^h \left(\sum_{t=0}^\infty \beta^t u^h\left(c_t^h\right)\right)$$

subject to

$$\sum_{h=1}^N c_t^h + \sum_{h=1}^N k_{t+1}^h = \sum_{h=1}^N k_t^h (1-\delta) + \sum_{h=1}^N a_t^h f^h\left(k_t^h\right),$$

where v^h is country h's welfare weight.
Productivity of country h follows the process

$$\ln a_{t+1}^h = \rho \ln a_t^h + \epsilon_{t+1}^h,$$

where $\epsilon_{t+1}^h \equiv \varsigma_{t+1} + \varsigma_{t+1}^h$ with $\varsigma_{t+1} \sim \mathcal{N}\left(0, \sigma^2\right)$ is identical for all countries and $\varsigma_{t+1}^h \sim \mathcal{N}\left(0, \sigma^2\right)$ is country-specific.

Table 2. Accuracy and speed in the multi-country model

	Polyn. degree	M1			Q(1)		
		Mean	Max	CPU	Mean	Max	CPU
N=2	1st	−4.09	−3.19	44 sec	−4.07	−3.19	45 sec
	2nd	−5.45	−4.51	2 min	−5.06	−4.41	1 min
	3rd	−6.51	−5.29	4 min	−5.17	−4.92	2 min
N=20	1st	−4.21	−3.29	20 min	−4.17	−3.28	3 min
	2nd	−5.08	−4.17	5 hours	−4.83	−4.10	32 min
N=40	1st	−4.23	−3.31	5 hours	−4.19	−3.29	2 hours
	2nd	−	−	-	−4.86	−4.48	24 hours
N=100	1st	−4.09	−3.24	10 hours	−4.06	−3.23	36 min
N=200	1st	−	−	-	−3.97	−3.20	2 hours

M1 means monomial integration with 2N nodes; Q(1) means quadrature integration with one node in each dimension; Mean and Max are mean and maximum unit-free Euler equation errors in log10 units, respectively; CPU is running time.

JEDC's (2011) special issue: a comparison of six methods.

30 Multi-country models with up to 10 countries:

Accuracy on a stochastic simulation:

- 1-st order perturbation method of Kollmann, Kim and Kim (2011): *max error = 6.310%*
- 2-nd order perturbation method of Kollmann, Kim and Kim (2011): *max error = 1.349%*
- Stochastic simulation algorithm of Maliar, Maliar and Judd (2011): *max error = 0.145%*
- Cluster-grid algorithms of Maliar, Maliar and Judd (2011): *max error = 0.009%*
- Smolyak's collocation method of Malin, Krueger and Kubler (2011): *max error = 0.030%*
- Monomial rule Galerkin method of Pichler (2011): *max error = 0.115%*

A new Keynesian model

- *Households* choose consumption and labor.
- Perfectly competitive *final-good firms* produce goods using intermediate goods.
- Monopolistic *intermediate-good firms* produce goods using labor and are subject to sticky price (á la Calvo, 1983).
- *Monetary authority* obeys a Taylor rule with zero lower bound (ZLB).
- *Government* finances a stochastic stream of public consumption by levying lump-sum taxes and by issuing nominal debt.
- *Six exogenous shocks:*
  (i) preference shock that scales the overall momentary utility e_t^u
  (ii) preference shock that affects marginal disutility of labor e_t^L
  (iii) premium in the return to bonds e_t^B
  (iv) shock to productivity of intermediate-good firms e_t^a
  (v) monetary-policy shock e_t^R
  (vi) government-spending shock e_t^G.

Equilibrium conditions

- *FOCs of the intermediate-good firms*

$$S_t = \frac{1}{\exp\left(e_t^a\right)} \cdot \exp\left(e_t^u + e_t^L\right) L_t^{\varphi} Y_t + \beta\theta E_t \left\{\pi_{t+1}^{\varepsilon} S_{t+1}\right\}$$

$$F_t = C_t^{-\gamma} Y_t + \beta\theta E_t \left\{\pi_{t+1}^{\varepsilon-1} F_{t+1}\right\}$$

$$\frac{S_t}{F_t} = \left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}$$

where β = discount factor; S_t and F_t = some constructed variables; θ = fraction of intermediate-good firms that cannot change price; $\varepsilon \geq 1$ = elasticity of substitution across different intermediate goods; C_t and Y_t = consumption and output; γ and φ = inverse of intertemporal elasticity of substitution of consumption and labor supply; π_t = gross inflation rate between $t-1$ and t.

Equilibrium conditions

- Law of motion for the price distortion D_t

$$D_t = \left[(1-\theta) \left[\frac{1-\theta \pi_t^{\varepsilon-1}}{1-\theta} \right]^{\frac{\varepsilon}{\varepsilon-1}} + \theta \frac{\pi_t^{\varepsilon}}{D_{t-1}} \right]^{-1}$$

- Euler equation from the household's problem

$$\exp\left(e_t^u\right) C_t^{-\gamma} = \beta \exp\left(e_t^B\right) R_t E_t \left[\frac{\exp\left(e_{t+1}^u\right) C_{t+1}^{-\gamma}}{\pi_{t+1}} \right]$$

where R_t is the gross nominal interest rate.

- Aggregate production

$$Y_t = \exp\left(e_t^a\right) L_t D_t$$

- Aggregate resource constraint

$$C_t + G_t = Y_t$$

where $G_t = \frac{\overline{G}}{\exp\left(e_t^G\right)} Y_t$ is government spending.

Equilibrium conditions

- Taylor rule with ZLB on the net nominal interest rate

$$R_t = \max\left\{1, \quad R_* \left(\frac{R_{t-1}}{R_*}\right)^{\mu} \left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_{N,t}}\right)^{\phi_y}\right]^{1-\mu} \exp\left(e_t^R\right)\right\}$$

where R_* is the long-run gross nominal interest rate; π_* is the inflation target; $Y_{N,t}$ is the natural level of output.

- Stochastic processes for shocks

$$\begin{aligned}
e_t^u &= \rho^u e_{t-1}^u + u_t^u, & u_t^u &\sim \mathcal{N}\left(0, \sigma_u^2\right) \\
e_t^L &= \rho^L e_{t-1}^L + u_t^L, & u_t^L &\sim \mathcal{N}\left(0, \sigma_L^2\right) \\
e_t^B &= \rho^B e_{t-1}^B + u_t^B, & u_t^B &\sim \mathcal{N}\left(0, \sigma_B^2\right) \\
e_t^a &= \rho^a e_{t-1}^a + u_t^a, & u_t^a &\sim \mathcal{N}\left(0, \sigma_a^2\right) \\
e_t^R &= \rho^R e_{t-1}^R + u_t^R, & u_t^R &\sim \mathcal{N}\left(0, \sigma_R^2\right) \\
e_t^G &= \rho^G e_{t-1}^G + u_t^G, & u_t^G &\sim \mathcal{N}\left(0, \sigma_G^2\right)
\end{aligned}$$

- 8 equations & 8 unknowns. 2 endogenous state variables, D_{t-1}, R_{t-1}.

Parameter values

We calibrate the model using the results in Smets and Wouters (2003, 2007), and Del Negro, Smets and Wouters (2007).

- Preferences: $\gamma = 1$; $\varphi = 2.09$; $\beta = 0.99$
- Intermediate-good production: $\varepsilon = 4.45$
- Fraction of firms that cannot change price: $\theta = 0.83$
- Taylor rule: $\phi_y = 0.07$; $\phi_\pi = 2.21$; $\mu = 0.82$
- Inflation target: $\pi_* \in \{1, 1.0598\}$
- Government to output ratio: $\overline{G} = 0.23$
- Stochastic processes for shocks:
  $\rho^u = 0.92$; $\rho^L = 0.881$; $\rho^B = 0.23$; $\rho^a = 0.2$; $\rho^R = 0.15$; $\rho^G = 0.95$
  $\sigma_u = 0.0054$; $\sigma_L = 0.006$; $\sigma_B = 0.0022$; $\sigma_a = 0.0082$; $\sigma_R = 0.0024$;
  $\sigma_G = 0.0038$

We compute 1st and 2nd order perturbation solutions using Dynare, and we compute 2nd and 3rd degree CGA solutions.

Table 3. Accuracy and speed in the new Keynesian model

	0% inflation target				0% inflation target and ZLB			
	PER1	PER2	CGA2	CGA3	PER1	PER2	CGA2	CGA3
CPU		9	363	664		9	445	914
Mean	-3.05	-3.81	-4.15	-4.26	-2.99	-3.40	-3.98	-4.05
Max	-0.89	-1.75	-1.85	-3.14	-0.90	-1.05	-1.93	-2.06
R_{min}	0.983	0.981	0.980	0.980	1.0	1.0	1.0	1.0
$R_{\max}$	1.040	1.038	1.039	1.038	1.040	1.038	1.039	1.039
$Fr_{(R\leq1)}$	8.20	8.13	8.27	8.46	6.78	6.66	8.63	8.34
$\triangle R$	0.23	0.05	0.11	–	0.90	0.94	0.14	–
$\triangle C$	1.35	0.18	0.17	–	3.22	3.58	1.06	–
$\triangle Y$	1.36	0.18	0.17	–	3.25	3.59	1.06	–
$\triangle L$	3.22	0.14	0.24	–	4.66	3.61	1.04	–
$\triangle \pi$	0.56	0.06	0.21	–	0.98	0.86	0.19	–

PER 1 and PER 2 = 1st and 2nd order Dynare solutions; CGA2 and CGA3 = 2nd and 3rd degree CGA; Mean and Max = average and maximum absolute errors (in log10 units); R_{min} and $R_{\max}$ = minimum and maximum R; $Freq_{(R\leq1)}$ = number of periods in which $R \leq 1$ (in %); $\triangle X$ = max difference from CGA3.

A stochastic simulation of time series solution for a new Keynesian economy

Figure 4a. A time-series solution to a new Keynesian model without ZLB

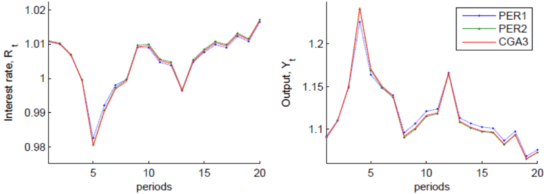
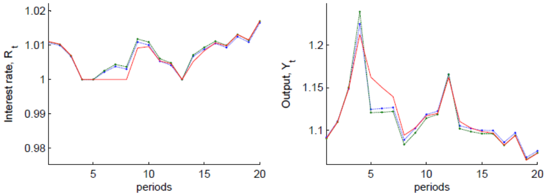


Figure 4b. A time-series solution to a new Keynesian model with ZLB



Conclusion

- CGA accurately solves models that were considered to be unfeasible until now.
- A mix of techniques taken together allows us to address the challenges of high-dimensional problems:
    - cluster-grid domain - a tiny fraction of the standard hypercube domain;
    - monomial and one-node integration rules;
    - fixed-point iteration for finding policy functions;
    - iteration-on-allocation and precomputation approaches for solving for intratemporal choice.
- A proper coordination of the above techniques is crucial for accuracy and speed.
- Parallelization and supercomputer (Condor).