Solving, Estimating, Selecting Nonlinear Dynamic Economic Models without the Curse of Dimensionality

Viktor Winschel

University of Mannheim
winschel@rumms.uni-mannheim.de

UNIVERSITÄT MANNHEIM

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Bonn-INS
2006

1 Introduction

Modern microfounded macromodels are

- *non-linear,*
- *with dynamic first order conditions*
- *and rational expectations*
- *and only partly observable variables.*

How to solve, estimate and select such models?

Solution: Judd '98, Krüger, Kübler '04, Smolyak '63
Estimation: Villaverde et.al. '04/'05
Selection: Gelfand, Dey '94, Geweke '98

Problems

Solution: Policy functions and rational expectations
Kronecker operator = curse of dimensionality
⇒ Smolyak operator = polynomial costs

Likelihood: Evaluation
SIR filter = Monte Carlo integration
⇒ Gaussian quadrature = Smolyak Kalman filter

Posterior: Density estimation
Metropolis-Hastings algorithm
⇒ Parallel Metropolis-Hastings algorithm

Model selection
Marginal likelihood

Relation to Literature

Towards applied nonlinear structural econometrics...

Constributions

	Villaverde et.al.	Winschel
Operator	Kronecker	Smolyak
Solution	Finite Elements	Smolyak Spectral
Rational Exp.	Gaussian	Smolyak Gaussian
Likelihood/Filter	SIR	Smolyak Kalman
Posterior	MH	Parallel MH
Generality	-	$f(s,x,z)$
Time	80h (Fortran)	5h (Ox)

$\Rightarrow$ Ox Code for a general model class $f(s,x,z)$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

2 Example

Macro model

$$
\begin{aligned}
\max_{\{c_t, l_t\}_{t=0}^{\infty}} U &= \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{(c_t^{\theta}(1-l_t)^{1-\theta})^{1-\tau}}{1-\tau} \\
\text{s.t.} & \\
y_t &= c_t + i_t \\
y_t &= e^{a_t} k_t^{\alpha} l_t^{1-\alpha} \\
k_{t+1} &= i_t + (1-\delta) k_t \\
a_{t+1} &= \rho a_t + \varepsilon_{t+1} \qquad \text{where} \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon).
\end{aligned}
$$

state: $s_t = \{k_t, a_t\}$
policies: $x_t = \{l_t, (c_t)\}$
shocks: $e_t = \{\varepsilon_t\}$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

First Order Conditions

Euler equation, intertemporal

$$\frac{(c_t^{\theta}(1-l_t)^{1-\theta})^{1-\tau}}{c_t} -$$

$$\beta \mathrm{E}_t \left(\frac{(c_{t+1}^{\theta}(1-l_{t+1})^{1-\theta})^{1-\tau}}{c_{t+1}} \left(1 - \delta + \alpha \frac{e^{a_{t+1}} k_{t+1}^{\alpha} l_{t+1}^{1-\alpha}}{k_{t+1}} \right) \right) = 0$$

labor/consumption trade off, intratemporal

$$\frac{1-\theta}{\theta} \frac{c_t}{1-l_t} - (1-\alpha) \frac{e^{a_t} k_t^{\alpha} l_t^{1-\alpha}}{l_t} = 0$$

capital transition

$$k_{t+1} = e^{a_t} k_t^{\alpha} l_t^{1-\alpha} - c_t + (1-\delta) k_t$$

productivity transition

$$a_{t+1} = \rho a_t + \varepsilon_t$$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

General Model Structure (Miranda/Fackler)

$$f(s, x, \mathrm{E}h(s, x, e', s', x')) = 0$$
$$s' = g(s, x, e')$$

Specified Functions

f	equilibrium conditions
h	expectation functions
g	state transitions

Variables

s	state variables
$x \in [a(s), b(s)]$	policy variables
z	expectation variables
e	stochastic shocks

Approximation Strategy

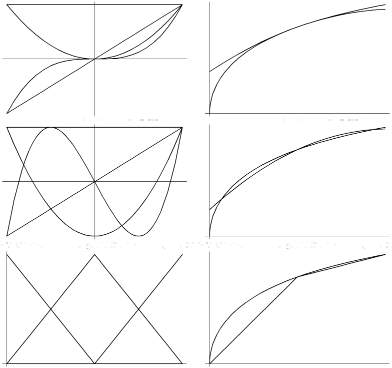
$x = \Psi(s) = \sum_{i=1}^{n} c_i \psi_i(s)$ policy functions

$$f(s, \Psi(s), \sum_j w_j h(s, \Psi(s), e'_j, s'_j, \Psi(s'_j))) = 0$$
$$s'_j = g(s, \Psi(s), e'_j)$$

UNIVERSITÄT MANNHEIM
Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

3 Solution

$f(x) = \sqrt{\log(x)}$ for $x \in [1,3]$



Perturbation

Spectral

Finite Elements

UNIVERSITÄT MANNHEIM

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Approximation

- Perturbation
- Spectral
- Finite Elements

– implicit function theorem, Taylor series of f, h, g

– matrix Riccati equation $\Rightarrow$ QZ decomposition

– higher order: one additional matrix inversion

pro: fast, no curse of dimensionality
con: deterministic steady state, inaccurate, singularities $\leftrightarrow$ local, no inequalities, symbolic or automated differentiation

UNIVERSITÄT MANNHEIM
Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Approximation

- Perturbation
- Spectral
- Finite Elements

– orthogonal basis polynomials

– non-zero at almost all points

pro: accurate, global
con: no inequalities, slow, start values

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Approximation

- Perturbation
- Spectral
- Finite Elements

– local basis functions

– zero at almost all points

pro: accurate, global/local $\leftrightarrow$ inequalities
con: slow, start values

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Integration

Monte Carlo Integration

$E(f(x)) = \int_{-\infty}^{\infty} f(x)p(x)dx \approx \sum_{i=1}^{N} f(x_i)/N$ for $x_i \sim p(x)$

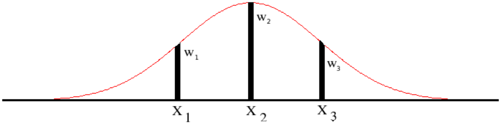
$\Rightarrow$ naive nodes x_i and weights $w_i = 1/N$

Gaussian quadrature

Discretize continous random variable

$E(f(x)) \approx \sum_{i=1}^{N} w_i f(x_i)$

$\Rightarrow$ optimal nodes x_i and weights w_i



Suffer from the curse of dimensionality

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Multidimensional Approximation

Kronecker Operator

$\{1,x,x^2\} \otimes \{1,y,y^2\} = \{1,x,y,xy,x^2,y^2,x^2y,xy^2,x^2y^2\}$

$\Rightarrow$ Problem: Curse of dimensionality, exponential costs

Complete polynomials in Taylor series

$F(x,y) \approx \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_3 x^2 + \alpha_4 y^2 + \alpha_5 xy$

$\Rightarrow$ Polynomial costs

$\Rightarrow$ Smolyak operator

Smolyak Operator

Approximation operator for univariate function/integrand

$$U^i(f) = \sum_{j=1}^{m_i} a_j^i f(x_j^i)$$

Kronecker operator $\otimes$ for $f : \mathbb{R}^d \to \mathbb{R}$

$$(U^{i_1} \otimes \cdots \otimes U^{i_d})(f) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_d=1}^{m_{i_d}} (a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_d}^{i_d}) f(x_{j_1}^{i_1}, \ldots, x_{j_d}^{i_d})$$

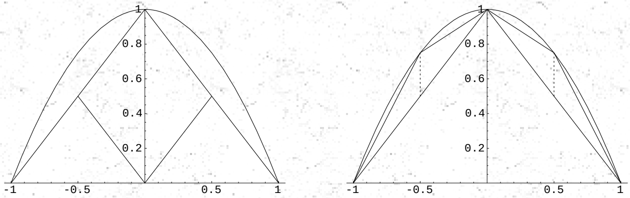
Smolyak operator with $\Delta^i = U^i - U^{i-1}$, $|\mathbf{i}| = i_1 + \ldots + i_d$

$$A_{q,d}(f) = \sum_{q-d+1 \le |\mathbf{i}| \le q} (-1)^{q-|\mathbf{i}|} \binom{d-1}{q-|\mathbf{i}|} (U^{i_1} \otimes \cdots \otimes U^{i_d})$$

$$= \sum_{|\mathbf{i}| \le q} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_d})(f) = A_{q-1,d}(f) + \sum_{|\mathbf{i}|=q} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_d})(f)$$

Hierarchical Surpluses

Archimedes approximated $1 - x^2$ with linear splines



Dotted lines = hierarchical surpluses $\longrightarrow 0$

Extension: Simple adaptive and dimension adaptive

Iteration

Start values $c^{(0)}$

$$
\begin{aligned}
x &= \psi(s)c^{(k)} \\
s'_j &= g(s, x, e'_j) \\
x'_j &= \psi(s'_j)c^{(k)} \\
z &= \sum_j w_j h(s, x, e'_j, s'_j, x'_j)
\end{aligned}
$$

a) Function Iteration

$$
\begin{aligned}
\text{solve} \quad & x(s,z) \text{ from } f(s,x,z) = 0 \\
\tilde{c} &= \Psi^{-1}(s)x(s,z) \\
c^{(k+1)} &= (1-\alpha)c^{(k)} + \alpha\tilde{c}
\end{aligned}
$$

b) Root finder (Broyden)

$$
c^{(k+1)} = c^{(k)} - \alpha[\partial r(s, c^{(k)})/\partial c]^{-1} r(s, c^{(k)})
$$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

4 Likelihood

Solution $\Psi(s)$ implies the state space model:

state transition: unobserved

$$\begin{aligned} s_t &= g^*(s_{t-1}, \Psi(s_{t-1}), e_t) \\ &= g(s_{t-1}, e_t) \qquad \Leftrightarrow \qquad p(s_t|s_{t-1}) \end{aligned}$$

measurement equation: observed

$$y_t = m(s_t, e_t^y) \qquad \Leftrightarrow \qquad p(y_t|s_t)$$

UNIVERSITÄT MANNHEIM
Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Model State Space

state transitions: unobserved

$$
\begin{aligned}
k_t &= e^{a_{t-1}} k_{t-1}^{\alpha} \psi_l(k_{t-1}, a_{t-1})^{1-\alpha} \\
&\quad - \psi_c(k_{t-1}, a_{t-1}) + (1-\delta) k_{t-1} \\
a_t &= \rho a_{t-1} + \varepsilon_t
\end{aligned}
$$

measurement equations: observed

$$
\begin{aligned}
y_t &= e^{a_t} k_t^{\alpha} \psi_l(k_t, a_t)^{1-\alpha} + e_t^y \\
l_t &= \psi_l(k_t, a_t) + e_t^l \\
i_t &= e^{a_t} k_t^{\alpha} \psi_l(k_t, a_t)^{1-\alpha} - \psi_c(k_t, a_t) + e_t^i
\end{aligned}
$$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Likelihood

$$\log \mathcal{L}(\theta) \equiv l(\theta) = \log p(y_{1:T}|\theta) = \log \prod_{t=1}^{T} p(y_t|y_{1:t-1})$$

Needed: sequential update of state estimates

$$p(s_t|y_{1:t}) \qquad \text{for } t = 1, \ldots, T$$

Prediction step:

$$p(s_t|y_{1:t-1}) = \int p(s_t|s_{t-1})p(s_{t-1}|y_{1:t-1})ds_{t-1}$$

Update step:

$$p(s_t|y_{1:t}) = \frac{p(y_t|s_t)p(s_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}$$

UNIVERSITÄT MANNHEIM

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Full Screen

Close

Filters

Problem: integrals are difficult to calculate

Analytical solution

- Kalman Filter for Linear Gaussian State Space

Usual Filter

- Extended Kalman Filter
- Unscented Kalman Filter ⇒ Smolyak Kalman filter
- Sequential Importance Resampling Particle Filter

UNIVERSITÄT MANNHEIM
Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Smolyak Kalman Filter

Assume: $p(s_{t-1} | y_{1:t-1}) = \mathcal{N}(\hat{s}_{t-1|t-1}, P^{ss}_{t-1|t-1})$

Prediction step: $p(s_t | y_{1:t-1})$

$$= \int p(s_t|s_{t-1})\, p(s_{t-1} | y_{1:t-1})\, ds_{t-1} de_t \Rightarrow \hat{s}_{t|t-1}, P^{ss}_{t|t-1}$$

Update step: $p(s_t | y_{1:t}) = \mathcal{N}(\hat{s}_{t|t}, P^{ss}_{t|t})$

$$\begin{aligned} \hat{s}_{t|t} &= \hat{s}_{t|t-1} + K_t(y_t - \hat{y}_{t|t-1}) \\ P^{ss}_{t|t} &= P^{ss}_{t|t-1} - K_t P^{yy}_{t|t-1} K_t^T \end{aligned}$$

$p(y_t|s_t)$ quadrature $\Rightarrow \hat{y}_{t|t-1}, P^{yy}_{t|t-1}, P^{sy}_{t|t-1}, K_t = P^{sy}_{t|t-1}(P^{yy}_{t|t-1})^{-1}$

period likelihood $l_t = \sum_{i=1}^{N} w^{(i)} \mathcal{N}(y_t - y^{(i)}_{t|t-1}; \hat{y}_{t|t-1}, P^{yy}_{t|t-1})$.

UNIVERSITÄT MANNHEIM

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Full Screen
Close

5 Posterior

posterior of unobservables (parameters) θ

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

variable of interest (deterministic function) ω

$$p(\omega, y, \theta) = p(\omega|y, \theta)p(y|\theta)p(\theta)$$

unobservables θ can be marginalized out

$$p(\omega|y) = \int p(\omega|y, \theta)p(\theta|y)\,d\theta$$

for small sample distribution of the variable of interest ω which can be any test statistic

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Metropolis-Hastings

$$
\begin{aligned}
&\texttt{for}\ \ n = 1,\ldots,N\\
&\texttt{1.}\ \ \texttt{generate}\ \ \theta_n^* = \theta_{n-1} + \varepsilon, \varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)\\
&\texttt{2.}\ \ \theta_n = \begin{cases} \theta_n^* & \texttt{if}\ \ U(0,1) \leq \frac{p(y^0_{1:t}|\theta_n^*)p(\theta_n^*)}{p(y^0_{1:t}|\theta_{n-1})p(\theta_{n-1})} \\ \theta_{n-1} & \texttt{otherwise} \end{cases}
\end{aligned}
$$

Critical choices:

$\Sigma_\varepsilon \Leftrightarrow$ acceptance ratio $\approx 30\% \Rightarrow$ training sequences
$\theta_0 \Leftrightarrow$ burn-in period $\theta_0 - \theta_J \Rightarrow$ robustness check
$J \Leftrightarrow$ diagnostic tests $\Rightarrow$ biased with one sequence

$$
\begin{aligned}
&\texttt{for}\ \ n = 1,\ldots,N,\ \ m = 1,\ldots,M,\ \ \texttt{random}\ \ m_1, m_2\\
&\texttt{1.}\ \ \theta_{m,n}^* = \theta_{m,n-1} + \gamma(\theta_{m_1,n-1} - \theta_{m_2,n-1}) + \varepsilon\\
&\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon \ll \gamma)
\end{aligned}
$$

6 Selection

Marginal likelihood of a model $M_i \in M$

$$p(y|M_i) = \int p(y|\theta_{M_i}, M_i)\, p(\theta_{M_i}|M_i)\, d\theta_{M_i}$$

allows data to assign probabilities to model M_i,

$$p(M_i|y, M) = \frac{p(y|M_i)\, p(M_i)}{p(y|M)}$$

transform model priors into posterior odds ratio,

$$\frac{p(M_i|y)}{p(M_j|y)} = \frac{p(M_i)}{p(M_j)} \frac{p(y|M_i)}{p(y|M_j)}.$$

and selects nonnested and quasi-true models.

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

Full Screen
Close

Marginal Likelihood

Model's marginal likelihood

$$p(y|M_i) = \int_{\Theta_{M_i}} p(y|\theta_{M_i}, M_i)\, p(\theta_{M_i}|M_i)\, d\theta_{M_i}$$

For any pdf $h(\theta_{M_i}|M_i)$, Gelfand, Dey '94

$$\mathsf{E}_{p(\theta_{M_i}|y,M_i)}\left(\frac{h(\theta_{M_i}|M_i)}{p(y|\theta_{M_i}, M_i)\, p(\theta_{M_i}|M_i)}\right) = p(y|M_i)^{-1}$$

Marginal likelihood estimate, Geweke '98

$$\hat{p}(y|M_i) = \left(\frac{1}{N}\sum_{n=1}^{N} \frac{h(\theta_{n,M_i})}{p(y|\theta_{n,M_i}, M_i)\, p(\theta_{n,M_i}|M_i)}\right)^{-1}$$

Title
Introduction
Example
Solution
Likelihood
Posterior
Selection
Results

7 Results

Smolyak

countries	Smolyak $\log_{10} r^E$	Smolyak points	Kronecker $\log_{10} r^E$	Kronecker points
1	-5.4/-6.5	13	-5.9/-6.4	25
1	-6.5/-6.8	29	-6.5/-6.7	36
2	-4.0/-4.9	41	-2.6/-3.5	81
2	-5.1/-6.5	137	-4.7/-6.1	256
3	-3.6/-4.8	85	-2.4/-2.7	64

$\log_{10} r^E$ is max/mean Euler error
in 10,000 simulated periods

$\Rightarrow$ Smolyak reduces Kronecker grid size substantially, already for small models

Estimation

	True	mean		std.dev.	
		Nonlin	Lin	Nonlin	Lin
θ	0.35	0.356555	0.359832	4.47E-04	3.56E-03
ρ	0.95	0.950678	0.942740	8.57E-04	6.74E-03
τ	50.00	49.385802	62.910866	9.01E-01	2.07E+01
α	0.40	0.399076	0.408750	1.25E-03	9.20E-03
β	0.99	0.990152	0.988453	2.09E-04	1.79E-03
δ	0.02	0.019809	0.025163	2.62E-04	2.26E-03
σ_a	0.035	0.035053	0.031274	1.17E-04	2.11E-03
σ_y	0.000158	0.000160	0.000674	6.10E-06	4.58E-04
σ_l	0.0011	0.001080	0.001351	2.55E-05	9.24E-05
σ_i	0.000866	0.000877	0.000994	1.89E-05	2.15E-04

$\Rightarrow$ Good nonlinear estimates, Biased linear estimates

Selection

	true	1209.9	-143,651.0
	max	1213.2	1190.3
p	Δ	Nonlinear	Linear
.1	28.8	-111.6	-140.4
.2	28.8	-110.9	-139.7
.3	28.8	-110.5	-139.3
.4	28.8	-110.2	-139.0
.5	28.8	-110.0	-138.8
.6	29.0	-109.6	-138.6
.7	29.0	-109.4	-138.5
.8	29.1	-109.3	-138.3
.9	29.2	-109.0	-138.2

$\Rightarrow$ Nonlinear model is clearly detected from data

Conclusion

- Smolyak operator
  - very effective
  - revision of standard approaches
  - useful for nonparameterics
- Smolyak Kalman filter
  - very fast
  - accurate maybe only for growth model
- Parallel MH
  - simplifies covariance choice
  - better global maximization properties
  - allows parallel estimation

Future

- OLG models
- Inequality Constraints
- Adaptive Smolyak Finite Elements
- Graph Theory for Causality Analysis
- Smolyak Gaussian Sum Filter
- C++, Cluster Implementation
- NEOS/AMPL model compiler for max. likelihood
- Statistical Model Compiler in Prolog...in 10 years :-)