Solving Heterogeneous Agent Models with Nonconvex Optimization Problems: Linearization and Beyond

Michael Reiter

IHS, Vienna and NYU Abu Dhabi

PASC18 Conference, July 2, 2018
Overview

Introduction

Specification of HA Models

Solution Method: Linearization

State and Value Function Aggregation

Global approximations
Outline

Introduction

Specification of HA Models

Solution Method: Linearization

State and Value Function Aggregation

Global approximations
Example 1: Chang/Kim, AER 2007, 97(5), p. 1939

- Heterogeneous households: idiosyncratic labor productivity follows Markov process
- Indivisible labor: work either zero hours or fixed number of hours
- Otherwise no labor market friction
- Objective: study observed "labor wedge": discrepancy between wage and observed MRS between consumption and leisure.

Why does

\[
w = \frac{U_L}{U_C} = \eta \frac{C}{1 - L}
\]  

not hold over the business cycle?

**Technology:**
Production function:

\[
Y_t = F(L_t, K_t, \lambda_t) = \lambda_t L_t^\alpha K_t^{1-\alpha}
\]

where \( \lambda_t \) is Markov with transition probability distribution \( \pi_{\lambda} \).
Chang/Kim 2007, Value function

\[ V(a, x; \lambda, \phi) = \max_{a' \in A, h \in \{0, \bar{h}\}} \left\{ u(c, h) + \beta \mathbb{E} V(a', x'; \lambda', \phi') \right\} \quad (3) \]

s.t.

\[ c = w(\lambda, \phi)xh + (1 + r(\lambda, \phi))a - a' \quad (4) \]
\[ a' \geq \bar{a} \quad (5) \]
\[ \phi' = T(\lambda, \phi) \quad (6) \]

where

- \( a \) is household assets
- \( x \) is exogenous individual productivity process
- \( \lambda \) is aggregate TFP process
- \( \phi \) is cross-sectional distribution of agents over \((a, x)\).
Properties of solution

Labor supply is determined by extensive margin: work or not
  ▶ Reaction of switch point to change in environment
  ▶ Fraction of households at the margin

Quality of linear approximation will depend on smoothness of cross-sectional distribution around switch points.
CK model, distribution around switch points

(183, 1, 7)
(299, 1, 8)
(420, 1, 9)
(537, 1, 10)
(644, 1, 11)
Technical problem of HA models

High dimensionality of the state space

- Impossible to solve agent optimization problem exactly.
- Impossible to follow exactly the dynamics of high-dimensional aggregate systems.
Outline method

- Stationary state without aggregate shocks
- Linearization around stationary state
- Almost exact state and value function reduction of linearized model
- Global approximation method; research question: how useful is state aggregation in linearized model for global approximation
- Implemented in Julia
Computation: Discrete dynamic programming

- for efficiency: make end-of-period state the discrete choice
- function evaluations at discrete grid points;
- thorough one-dimensional search for global maximum
  - first evaluate at grid points (no interpolation necessary), find local maximum
  - between grid points:
    1. evaluate at 8 equidistant points (easy for SIMD)
    2. Newton method starting from local maximum

Parallelization

1. SIMD vectorization: needs SIMD versions of log etc. (in Intel MKL, not generally available now)
2. Across cores:
   - exogenous transition: matrix multiplication
   - endogenous transition: separate for each grid point; only needs continuation value for given exogenous state
Computation: Invariant distribution

Three possible methods

1. Time iteration (sparse matrix times vector multiplication); can probably be further accelerated

2. Sparse eigenvector problem (to largest eigenvalue, know to be 1)

3. Finding nullspace of $(\Pi - I)$.

First one is very robust and fastest in most cases.
Computation: Linarization

1. Automatic differentiation, forward mode (dense and sparse; written in Julia)
2. SVD for state and value function reduction (parallelized in OpenBlas library)
3. Dense matrix multiplications (parallelized)
4. Sparse matrix-vector multiplications
5. Solution of RE equations:
   5.1 QZ (generalized Schur decomposition): currently not parallelized
   5.2 Backward iteration: dense matrix multiplications, parallelized
Computation time in Chang/Kim model

Size of problem:
- 17000 state variables
- 6800 Elements of value vector
- 6800 Decisions

Computing time in seconds (laptop)
- steady state: 19 (not well parallelized)
- differentiation: 35
- value function reduction: 1
- state reduction: 16
- iterative solution linearized model: 10 (time iteration)
Results linearized model

Linearized model:

- Number of states reduced from 17000 to around 400
- Number of value function parameters reduced from 6800 to around 200
- Accuracy: aggregation error $\approx 10^{-12}$.
- Reduced states obtained from linearized model not useful for nonlinear solution
- Smooth state aggregation (aggregate mass at adjacent bins into one statistic) works better (easier to predict).
Preliminary results global approximation

- Compute temporary equilibrium, given next period’s continuation value
- Take continuation value from linearized model
- Simulate long series
- Test whether aggregate states
  - predict future prices (KL ratio, ”ResidKL”)
  - predict their own future values (”ResidOwn”)
  - predict future prices only with states that predict themselves well (”ResidKL2”)

<table>
<thead>
<tr>
<th></th>
<th>ResidKL</th>
<th>ResidOwn</th>
<th>Resid KL2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gramian</td>
<td>6.9(-6)</td>
<td>0.062</td>
<td>2.5(-4)</td>
</tr>
<tr>
<td>BalancedRed</td>
<td>2.2(-4)</td>
<td>0.066</td>
<td>6.0(-4)</td>
</tr>
<tr>
<td>Smooth</td>
<td>8.1(-6)</td>
<td>0.013</td>
<td>1.3(-5)</td>
</tr>
</tbody>
</table>
Outline

Introduction

Specification of HA Models

Solution Method: Linearization

State and Value Function Aggregation

Global approximations
Outline

Introduction

Specification of HA Models

Solution Method: Linearization

State and Value Function Aggregation

Global approximations
Stylized Representation of Heterogeneous Agent Models

- Continuum (mass 1) of agents, ex ante identical. Index an agent by subscript $h$.
- Individual endogenous state: $x_{h,t} \in \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N\}$.
- Individual exogenous shock: $y_{h,t} = \bar{y}_j$ with probability $p_j$, $j = 1, \ldots, J$.
  $y_{h,t}$ is i.i.d. across agents and time (so it is not a state, for ease of notation).
- Agents differ ex post because of individual history $y_{h,t}$.
- Aggregate exogenous process $Z_t$ (first order Markov).
- Decisions are taken after $Z_t$ and $y_{h,t}$ are realized.
Agents’ objective

\[
\max_{a_{h,t}} \mathbb{E} \sum_{t=0}^{\infty} \beta^t U(a_{h,t}, x_{h,t-1}, y_{h,t}, Z_t, X_{t-1}, A_t) \tag{7}
\]

subject to \(a_{h,t} \in G(x_{h,t-1})\) and

\[
\text{prob}(x_{h,t} = \bar{x}_j | x_{h,t-1} = \bar{x}_i) = P_{i,j}(a_{h,t}, x_{h,t-1}, y_{h,t}, Z_t, X_{t-1}, A_t) \tag{8}
\]

where

- \(\phi_{i,t}\) is mass of agents at state \(\bar{x}_i\) at end of period \(t\).
- \(X_t \equiv \sum_{i=1}^{N} \bar{x}_i \phi_{i,t}\) is average state.
- \(A_t \equiv \sum_{i=1}^{N} a(\bar{x}_i) \phi_{i,t}\) is average action.
- \(G(x_{h,t-1})\): set of feasible actions \(a_{h,t}\) conditional on \(x_{h,t-1}\).

Notice: the distribution \(\vec{\phi}\) directly affects utility (7) and the transition function (8) only through the aggregates \(X\) and \(A\)!
Structure of exact solution

- Aggregate State Space
  - $N$ aggregate state variables:
    1. $\phi_{i,t}$, $i = 1, \ldots, N$ with $N - 1$ degrees of freedom
    2. $Z_t$

- Individual decision function

\[
a_{h,t} = \alpha(x_{h,t-1}, y_{h,t-1}, \phi_{t-1}, Z_t)
\]
Aggregate Law of Motion

Given

- the individual decision function $\alpha(.)$
- last period’s state $\vec{\phi}_{t-1}$
- this period’s realizations $Z_t$ (realizations of $y_{h,t}$ do not matter because of law of large numbers)

we get

$$\phi_{m,t} = \sum_{n=1}^{N} \sum_{j=1}^{J} \phi_{n,t-1} \cdot p_j \cdot P_{n,m}(\alpha(\bar{x}_n, \bar{y}_j, \vec{\phi}_{t-1}, Z_t), \bar{x}_n, \bar{y}_j, Z_t, X_{t-1}, A_t)$$

(10)

Denote ALM by $\Phi(\vec{\phi}_{t-1}, Z_t)$. 
Notice that, for given decision function $\alpha(\bar{x}_n, \bar{y}_j, \bar{\phi}_{t-1}, Z_t)$, (10) is linear in $\bar{\phi}_{t-1}$, so we can write the ALM in matrix form as

$$\bar{\phi}_t = \Pi_t \bar{\phi}_{t-1}$$

where the elements of the transition matrix $\Pi$ are given by

$$\Pi_{m,n,t} = \sum_{j=1}^{J} p_j \cdot P_{n,m}(\alpha(\bar{x}_n, \bar{y}_j, \bar{\phi}_{t-1}, Z_t, A_t), \bar{x}_n, \bar{y}_j, Z_t, X_{t-1})$$
Recursive equilibrium

- Agents’ optimization: Bellman equation

\[
V(\bar{x}_n, \bar{\phi}_{t-1}, Z_t) = \max_a \left\{ \sum_j p_j \left[ U(a, \bar{x}_n, \bar{y}_j, Z_t, X_{t-1}, A_t) \right] + \beta \sum_{n=1}^{N} P_{n,m}(a, \bar{x}_n, \bar{y}_j, Z_t, X_{t-1}, A_t) E_t V(\bar{x}_m, \bar{\Phi}(\bar{\phi}_{t-1}, Z_t), Z_{t+1}) \right\}
\]

which gives optimal decision \( \alpha(\bar{x}_n, \bar{\phi}_{t-1}, Z_t) \).

- \( \bar{\Phi}(\bar{\phi}_{t-1}, Z_t) \), satisfies (10) given decision rule \( \alpha(x, \bar{\phi}, Z) \).

- \( y_i \) and \( Z \) follow exogenous processes

Notice: expectation \( E_t \) is over \( Z_{t+1} \).
Equation System for Linearization

\[ Z_t = \rho Z_{t-1} + \varepsilon_t \]  \hspace{1cm} (12)

FOCs (at points with interior solution) and values:

\[ 0 = \sum_j p_j \left[ U_a \left( a_n, t, \bar{y}_j, Z_t, X_{t-1} \right) \right. \]

\[ + \beta \sum_{m=1}^{N} \frac{\partial P_{n,m}(a_n, \bar{x}_n, \bar{y}_j, Z_t, X_{t-1})}{\partial a_{n,t}} E_t V_{t+1}(\bar{x}_m) \]  \hspace{1cm} (13)

\[ V_t(\bar{x}_n) = \sum_j p_j \left[ U \left( a_n, t, \bar{y}_j, Z_t, X_{t-1} \right) \right. \]

\[ + \beta \sum_{m=1}^{N} P_{n,m}(a_n, \bar{x}_n, \bar{y}_j, Z_t, X_{t-1}) E_t V_{t+1}(\bar{x}_m) \]  \hspace{1cm} (14)

Distribution dynamics (10) and definition

\[ X_t \equiv \sum_{i=1}^{N} \bar{x}_i \phi_{i,t} \]  \hspace{1cm} (15)
Equation System for Linearization, ctd.

- Equus. (10), (12), (13), (14) and (15) define a dynamic system of $3N + 2$ nonlinear equations in $3N + 2$ variables in each period $t$:
  - $Z_t$
  - $a_{n,t}$, $n = 1, \ldots, N$
  - $\phi_{n,t}$, $n = 1, \ldots, N$
  - $X_t$
  - $V_t(\bar{x}_n)$, $n = 1, \ldots, N$

Notice: since $\sum_{i=1}^{N} \phi_i = 1$, one equation in (10) is redundant and will be replaced by $\sum_{i=1}^{N} \tilde{\phi}_i = 1$. 
Steady state without aggregate shocks

- Set all $Z_t$ (but not the $y_{h,t}!$) to their unconditional expectation $Z^*$.
- Then Equs. (10), (12), (13), (14) and (15) define a static system of $3N + 1$ nonlinear equations in the $3N + 2$ variables $Z^*$ and $(a^*_n, \phi^*_n, V^*(\bar{x}_n))$ for $n = 1, \ldots, N$.
- Define

$$X^* \equiv \sum_{i=1}^{N} \bar{x}_i \phi^*_i \quad (16)$$

The solution can be computed by standard methods:
1. Guess $X^*$
2. Given $X^*$, compute the HH solution (by DP) and the corresponding distribution $\phi^*$.
3. Iterate over $X^*$ until (16) is satisfied.
Linearization around steady state (Reiter 2009)

- Stack all variables into the vector

\[
\theta_t = \begin{bmatrix}
Z_t \\
X_t \\
a_{n,t}, n = 1, \ldots, N \\
\phi_{n,t}, n = 1, \ldots, N \\
V_t(\bar{x}_n), n = 1, \ldots, N
\end{bmatrix}
\]

- The linearized model takes the form

\[
\Lambda \theta_{t-1} + \Gamma \theta_t + E_t \Phi \theta_{t+1} + \Psi \epsilon_t = 0 \quad (17)
\]

- To obtain (17), differentiate the $3N + 2$ equations (10), (12), (13), (14), (15) w.r.t. all elements of $\theta_{t-1}, \theta_t, \theta_{t+1}$, at the deterministic steady state.
Split equation system

Organize the linear equation system as

\[
\begin{bmatrix}
\Lambda_{ss} & 0 & 0 \\
\Lambda_{ys} & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
{s_{t-1}} \\
y_{t-1} \\
v_{t-1}
\end{bmatrix}
+ \begin{bmatrix}
\Gamma_{ss} & \Gamma_{sy} & \Gamma_{sv} \\
\Gamma_{ys} & \Gamma_{yy} & \Gamma_{yv} \\
\Gamma_{vs} & \Gamma_{vy} & \Gamma_{vv}
\end{bmatrix}
\begin{bmatrix}
{s_{t}} \\
y_{t} \\
v_{t}
\end{bmatrix}
+ \begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
\Phi_{sv} \\
\Phi_{yv} \\
\Phi_{vv}
\end{bmatrix}
\begin{bmatrix}
{s_{t+1}} \\
y_{t+1} \\
v_{t+1}
\end{bmatrix}
+ \begin{bmatrix}
\Psi_{s} \\
\Psi_{y} \\
0
\end{bmatrix}
\epsilon_{t} = 0 \quad (18)
\]

- Only \(v\) appears with time index \(t + 1\)
- Only \(s\) appears with time index \(t - 1\)
- Only \(y\), not \(s\) enters equations for \(v\).
- We assume \(\Gamma_{ss}, \Gamma_{yy}, \Gamma_{vv}\) are regular.
- We assume \(\Gamma_{ss} \) and \(\Gamma_{vv}\) are easy to invert.
How to split

- \( s \) are all the variables that appear with time index \( t - 1 \).
- \( v \) are all the variables that appear with time index \( t + 1 \).
- \( y \) are all the other variables.

Split the equation systems into equations so as to satisfy the constraints implicit in (18). This cannot be fully automatized, therefore: name each state transition equation equal to the name of the state variable.
Normalized equation system

Rewrite (18) as

\[
\begin{bmatrix}
\Gamma_{ss}^{-1} \Lambda_{ss} \\
\Lambda_{ys} \\
0
\end{bmatrix}
\begin{bmatrix}
s_t^{-1}
\end{bmatrix}
+ \begin{bmatrix}
I & \Gamma_{ss}^{-1} \Gamma_{sy} & \Gamma_{ss}^{-1} \Gamma_{sv} \\
\Gamma_{ys} & \Gamma_{yy} & \Gamma_{yv} \\
\Gamma_{vs}^{-1} \Gamma_{vv} & \Gamma_{vv}^{-1} \Gamma_{vy} & I
\end{bmatrix}
\begin{bmatrix}
s_t \\
y_t \\
v_t
\end{bmatrix}
+
E_t
\begin{bmatrix}
\Gamma_{ss}^{-1} \Phi_{sv} \\
\Phi_{yy} \\
\Gamma_{vv}^{-1} \Phi_{vv}
\end{bmatrix}
\begin{bmatrix}
v_{t+1}
\end{bmatrix}
+ \begin{bmatrix}
\Gamma_{ss}^{-1} \Psi_s \\
\Psi_y \\
0
\end{bmatrix}
\epsilon_t = 0
\]
Model Reduction (Aggregation)

- State reduction: choose selection matrix $\tilde{M}$ that computes statistics ("moments") of the cross-sectional distribution

\[ m_t = \tilde{M}s_t \quad (20) \]

$\text{dimen}(m) << \text{dimen}(s)$.

- Value reduction: choose matrix $\bar{V}$ with $\bar{V}'\bar{V} = I$ that spans the space in which the value function lies:

\[ v_t = \bar{V}f_t \quad (21) \]

$\text{dimen}(f) << \text{dimen}(v)$. 

Almost Exact Value Function Reduction

Forward looking equation system (simplified notation)

\[ v_t = -E_t \left( \Gamma_{vy} y_t + \Phi_{vv} v_{t+1} \right) \]  \hfill (22)

Iterating gives

\[ v_t = -E_t \left[ \Gamma_{vy} y_t + \Phi_{vv} \Gamma_{vy} y_{t+1} + \Phi_{vv}^y \Gamma_{vy} y_{t+y} + \ldots \right] \]

\[ = -E_t \sum_{i=0}^{\infty} \Phi_{vv}^i \Gamma_{vy} y_{t+i} \]  \hfill (23)

This implies that \( v_t \) is spanned by the \( \Phi_{vv}^i \Gamma_{vy} \).

For \( \bar{V} \), choose orthonormal basis of \( \bigcup_{i=0}^{k} \Phi_{vv}^i \Gamma_{vy} \) for some finite \( k \) (high accuracy because of discounting)

Essential: \( \Gamma_{vy} \) must have small rank!
State reduction (aggregation)

- Choose an \( n_m \times n_s \) matrix \( \tilde{M} \) with \( n_m < n_s \) and define

\[
m_t = \tilde{M}s_t
\]  

(24)

Interpretation: \( m_t \) denotes the statistics ("m" is a memo of "moments) of the cross-sectional distribution that agents based their decision on (bounded rationality). We assume that those statistics are linear functions of the distribution.

- The matrix should be such that there exist \( \tilde{\Lambda}_{ys} \) and \( \tilde{\Gamma}_{ys} \) with

\[
\Lambda_{ys} = \tilde{\Lambda}_{ys} \tilde{M}, \quad \Gamma_{ys} = \tilde{\Gamma}_{ys} \tilde{M}, \quad \Gamma_{vs} = \tilde{\Gamma}_{vs} \tilde{M}
\]  

(25)

(25) is satisfied if the rows of \( \Lambda_{ys} \) and \( \Gamma_{ys} \) are spanned by the rows of \( \tilde{M} \):

\[
[\Lambda'_{ys} \quad \Gamma'_{ys} \quad \Gamma'_{vs}] \in span(\tilde{M}')
\]  

(26)

Using (25) in the second block of equations in (18), the terms in \( s \) are replaced by terms in \( m \).
Exact State Reduction: Gramians

Consider an $n \times n$ matrix $A$ and an $m \times n$ matrix $C$ and compute the $m \cdot n \times n$ observability matrix

$$Q = \begin{bmatrix}
C \\
CA \\
CA^2 \\
\vdots \\
CA^{n-1}
\end{bmatrix}$$

(27)

Define $k \leq n$ as the rank of $Q$. The SVD of $Q$ can be written as

$$Q = [U_1 \quad U_2] \begin{bmatrix}
S & 0 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
V'_1 \\
V'_2
\end{bmatrix} = U_1 SV'_1$$

(28)

$$S \equiv diag(\sigma_1, \ldots, \sigma_k)$$

(29)

where $U_1$ and $V_1$ have dimension $m \cdot n \times k$, $U'_1 U_1 = V'_1 V_1 = I_k$. 
Computing an $\tilde{M}$ with $\tilde{MA} = \hat{A}\tilde{M}$

From the Cayley-Hamilton theorem, there exists a $\Lambda$ such that

$$QA = \Lambda Q$$  \hspace{1cm} (30)

Using (28) in (30) we get $U_1 SV_1' A = \Lambda U_1 SV_1'$. Premultiplying by $S^{-1} U_1'$ we get $V_1' A = S^{-1} U_1' \Lambda U_1 SV_1'$ (notice that $S$ is invertible). Setting $\tilde{M} = V_1'$ we get

1. $\tilde{M}\tilde{M}' = I_k$
2. $\tilde{M}$ can be interchanged with $A$:

$$\tilde{MA} = \hat{A}\tilde{M}$$  \hspace{1cm} (31)

with $\hat{A} = S^{-1} U_1' \Lambda U_1 S$ being a $k \times k$-matrix. From 1. it follows that $\hat{A} = \tilde{MA}\tilde{M}'$.

3. We have $C' \in \text{span}(\tilde{M}')$:

$$C = \begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix} Q = \begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix} U_1 S \tilde{M}$$  \hspace{1cm} (32)
Almost exact state aggregation (Reiter 2010a)

Assume that $Q$ has rank $k << n$ when choosing $A = \Gamma_{ss}^{-1}\Lambda_{ss}$ and $C$ such that

$$
\begin{bmatrix}
\Lambda'_{ys} & \Gamma'_{ys} & \Gamma'_{vs}
\end{bmatrix} \in \text{span}(C) \quad (33)
$$

From the result above, there is a $k \times n$-matrix $\tilde{M}$ such that

- $\Lambda_{ys} = \tilde{\Lambda}_{ys} \tilde{M}$, $\Gamma_{ys} = \tilde{\Gamma}_{ys} \tilde{M}$, $\Gamma_{vs} = \tilde{\Gamma}_{vs} \tilde{M}$
- There exists an $n_m \times n_m$-matrix $\hat{T}$ such that

$$
\tilde{M}\Gamma_{ss}^{-1}\Lambda_{ss} = \hat{T}\tilde{M} \quad (34)
$$

Then $\tilde{M}\Gamma_{ss}^{-1}\Lambda_{ss}s_t = \hat{T}m_t$. 
Normalized and reduced equation system

Combining almost exact state and value reduction we have

1. \( v_t = \bar{V}' f_t \) with \( \bar{V}' \bar{V} = I \)

2. \( \Lambda_{ys} = \tilde{\Lambda}_{ys} \bar{M}, \quad \Gamma_{ys} = \tilde{\Gamma}_{ys} \bar{M}, \quad \Gamma_{vs} = \tilde{\Gamma}_{vs} \bar{M} \)

3. \( \bar{M} \Gamma_{ss}^{-1} \Lambda_{ss} s_t = \hat{T} m_t \)

We can transform (19) to the much smaller system

\[
\begin{bmatrix}
\hat{T} \\
\tilde{\Lambda}_{ys} \\
0
\end{bmatrix} m_{t-1} + \begin{bmatrix}
I & \bar{M} \Gamma_{ss}^{-1} \Gamma_{sy} & \bar{M} \Gamma_{ss}^{-1} \Gamma_{sv} \bar{V} \\
\tilde{\Gamma}_{ys} & \Gamma_{yy} & \tilde{\Gamma}_{yv} \bar{V} \\
\Gamma_{vv}^{-1} \tilde{\Gamma}_{vs} & \Gamma_{vv}^{-1} \Gamma_{vy} & I
\end{bmatrix}
\begin{bmatrix}
m_t \\
y_t \\
f_t
\end{bmatrix} + 
E_t \begin{bmatrix}
\bar{M} \Gamma_{ss}^{-1} \Phi_{sv} \bar{V} \\
\Phi_{yv} \bar{V} \\
\Gamma_{vv}^{-1} \Phi_{vv} \bar{V}
\end{bmatrix} f_{t+1} + 
\begin{bmatrix}
\bar{M} \Gamma_{ss}^{-1} \Psi_s \\
\Psi_y \\
0
\end{bmatrix} \epsilon_t = 0 \quad (35)
Asymmetry state reduction versus value reduction

1. Iterating the Bellman equation forward we can compute a basis for the subspace in which the value function lives.

2. The subspace in which the cross-sectional distribution lives depends on the shocks. It cannot be determined without solving the model.
   The commutation property (34) says that, knowing the statistics $m_t$, the exact distribution does not matter for the solution. It does not give a good proxy distribution. Notice that the optimal state reduction does not depend on the shocks!
State reduction Chang/Kim model

Linearized model:
- Number of states reduced from 17000 to around 400
- Number of value function parameters reduced from 6800 to around 200
- Accuracy: aggregation error $\approx 10^{-12}$. 
Outline

Introduction

Specification of HA Models

Solution Method: Linearization

State and Value Function Aggregation

Global approximations
A class of models for global approximations

- State variables at time $t$:
  - Individual endogenous state $x_{t-1}$
  - Individual exogenous state $\xi_t$ (Markov chain)
  - End-of-period cross-sectional distribution over individual states $\phi_t$ (Markov chain)
  - Aggregate exogenous state $z_t$ (Markov chain)

- Individual and equilibrium variables:
  - Individual decision $x$; optimal decision function $\mathcal{X}(x, \xi, \phi, z)$ solving
    \[
    \max_{x_t, x_{t+1}, \ldots} \sum \beta^i E_t U(x_{t+i-1}, \xi_{t+i}, P_{t+i})
    \]  
    \[ (36) \]
  - Aggregate variables
    \[
    P = \mathcal{P}(\phi_t, z_t, \mathcal{X})
    \]  
    \[ (37) \]

- Aggregate law of motion for distribution
  \[
  \phi_t = \mathcal{T}(\phi_{t-1}, z_{t+1}, \mathcal{X})
  \]  
  \[ (38) \]
Interpolating the value function

Model the exogenous driving force $z$ as finite Markov chain, so that we do not interpolate in $z$.
Interpolation is a 3-step process:

1. For each possible realization of the exogenous aggregate state $z$, there is a separate value function. No interpolation in $z$. This should allow to handle large shocks.

2. For each value of $\xi$ and $z$, beginning-of-period value function $V$ is (linearly) interpolated in moments $m$ to obtain end-of-period value function $\tilde{V}_t (\bar{x}_i, \bar{\xi}_j, m, \bar{z}_l)$ at grid points.

3. The endogenous continuous state $x$ is piecewise quadratically interpolated. The quadratic components for each interval are obtained from the steady state. This means, the second derivative in $x$ does not depend on $z$ and $m$. 
Interpolating the value function in aggregate states

In step 2., the beginning-of-period value function $V$ is (linearly) interpolated in moments $m$ to obtain end-of-period value function $\tilde{V}$:

$$
\tilde{V}_t (\bar{x}_i, \bar{\xi}_j, m, \bar{z}_l) = \sum_{j'=1}^{n_z} \sum_{l'=1}^{n_Z} pr[\bar{\xi}_{j'}|\bar{\xi}_j] pr[\bar{z}_{l'}|\bar{z}_l] \times

\left[ v_{i,j,l}^* + \sum_{k=1}^{n_m} v_{i,j,k,l}(m_k - m_k^*) \right] (39)

The value function is represented in the tensors $v_{i,j,l}^*$ and $v_{i,j,k,l}$. The latter has dimension $n_x \times n_z \times n_m \times n_Z$. 
Having defined $\tilde{V}_t(x, \bar{\xi}_j, m, \bar{z}_l)$ on grid points of $x$ by (39), we set

$$
\tilde{V}(x, \bar{\xi}_j, m, \bar{z}_l) = (1 - p) \tilde{V}(\bar{x}_I(x), \bar{\xi}_j, m, \bar{z}_l) + p(1 - p) \tilde{Q}(I(x), j)
$$

(40)

where

$$
p = \frac{x - \bar{x}_I(x)}{\bar{x}_I(x+1) - \bar{x}_I(x)}
$$

(41)
Simulating with temporary equilibrium

Given:

- end-of-period value function \( \tilde{V}_t (\bar{x}_i, \bar{\xi}_j, m, \bar{z}_l) \)
- Perceived law of motion from beginning-of-period moments to end-of-period moments: \( \tilde{m}_t(m_t) \).

Simulation, starting from arbitrary distribution \( \phi_0 \):

1. At beginning-of-period distribution \( \phi_{t-1} \), find equilibrium vector \( P \) such that \( P = P(\phi_{t-1}, z_t, X) \) where \( X \) is the decision function solving

\[
\max_{\xi_t} U(x_{t-1}, \xi_t, P_t) + \beta \tilde{V}(x, \xi_t, \tilde{m}_{k,l}, \bar{z}_l)
\]

2. Next period distribution \( \phi_t = T(\phi_{t-1}, z_t, X) \)
3. Iterate 1. and 2.
Starting the nonlinear solution algorithm

- Choose \( \bar{M} \) to give a vector of statistics \( m = \bar{M}s \), and compute \( \bar{P} \).
- Choose an aggregate grid \( (\bar{m}_1, \bar{m}_2, \ldots, \bar{m}_N) \), in deviations from (this grid may use the covariance matrix \( \Sigma_M \) of \( m \) obtained from the linearized model).
- For each aggregate grid point \( (k, l) \), guess a vector \( \tilde{m}_{k,l} \) of end-of-period aggregate moments.
- Define a vector of \( n_e \) equilibrium variables (such as prices) and \( n_e \) equations (such as market clearing).
- Initialize value function from linearized model.
Updating step of nonlinear solution algorithm

Given:
- Beginning of period value function, in form $v_{i,j,k,l}$.
- Transition beginning-end of period: $m_{k,l} \rightarrow \tilde{m}_{k,l}$

For each aggregate grid point $(\tilde{m}_k, \tilde{z}_l)$,
1. compute $\tilde{V}(:, :, \tilde{m}_k, l, \tilde{z}_l)$ by (39).
2. find equilibrium vector $P$ such that $P = \mathcal{P}(\phi_t, z_t, \lambda')$ where $\lambda'$ is the decision function solving

   $$\max_{\xi_t} U(x_{t-1}, \xi_t, P_t) + \beta \tilde{V}(x, \xi_t, \tilde{m}_k, l, \tilde{z}_l)$$

   (43)

3. Given equilibrium $p$, update the value function

   $$V(\tilde{x}_i, \tilde{\xi}_j, \tilde{m}_k, \tilde{z}_l) = U\left(x'_{i,j}, p, \tilde{x}_i, \tilde{\xi}_j, \tilde{m}_k, \tilde{z}_l\right) + \beta \tilde{V}(::, ::, \tilde{m}_k, l, \tilde{z}_l)$$

   (44)

4. Update $\tilde{m}_{k,l}$ by $\tilde{m}_{k,l} = \tilde{M}T\left(\bar{P}(\tilde{m}_k, l, \tilde{z}_l), z_t, \lambda\right)$
Initializing the value function

The beginning of period value function

\[ V(x_{t-1}, \xi_t, m_{t-1}, z_t) = V^* + \bar{V} \left[ A_{vm}(m_{t-1} - m^*) + B_v(z_t - z^*) \right] \]

(45)

Notice that we assume that the model is written such that each exogenous shock follows an independent AR process, such that the impact matrix \( B \) gives the derivative of the solution w.r.t. the exogenous driving force. [Easy to generalize.]


