Discrete State Dynamic Games

Kenneth L. Judd, Hoover Institution

May 13, 2020

Discrete-time Dynamic Games

- A discrete-time stochastic game with a finite number of states is often just called a "stochastic game"
  - Ericson-Pakes model of industry dynamics is an example
  - Pakes-Mcguire presents a computational method
- Definition of states and actions
  - State of the game in period t is $\omega_t \in \Omega$; finite number of states
  - N players.
  - Player i's action at t is $x_t^i \in \mathbb{X}^i(\omega_t)$, the set of feasible actions
  - The players' actions in period t is $x_t = \left(x_t^1, \ldots, x_t^N\right)$. As usual, x_t^{-i} denotes $\left(x_t^1, \ldots, x_t^{i-1}, x_t^{i+1}, \ldots, x_t^N\right)$.
- Apologies for change in notation. Here x_t^i denotes actions and ω_t^i denotes states

Dynamics and payoffs

- Dynamics
  - Changes in states are determined by a Markov process
  - Law of motion is

$$\Pr\left(\omega'|\omega_t, x_t\right) = \prod_{i=1}^{N} \Pr^i\left(\left(\omega'\right)^i|\omega_t^i, x_t^i\right),$$

  where $\Pr^i\left(\left(\omega'\right)^i|\omega_t^i, x_t^i\right)$ is the transition probability for player i's state.
- Payoff
  - Player i receives $\pi^i(x_t, \omega_t)$ when players' actions are x_t and the state is ω_t.
  - At the beginning of the next period player i receives a payoff $\Phi^i\left(x_t, \omega_t, \omega_{t+1}\right)$ IF there is a change in the state. For example, I may order a machine to come tomorrow but perhaps it does not.

Nash equilibrium

- Bellman equation for player i is

$$V^i(\omega) = \max_{x^i} \pi^i\left(x^i, X^{-i}(\omega), \omega\right) + \\ \beta E_{\omega'}\left\{\Phi^i\left(x^i, X^{-i}(\omega), \omega, \omega'\right) + V^i(\omega') \,|\omega, x^i, X^{-i}(\omega)\right\}$$

- Player strategy is

$$X^i(\omega) = \arg\max_{x^i} \pi^i\left(x^i, X^{-i}(\omega), \omega\right) + \\ \beta E_{\omega'}\left\{\Phi^i\left(x^i, X^{-i}(\omega), \omega, \omega'\right) + V^i(\omega') \,|\omega, x^i, X^{-i}(\omega)\right\}$$

- Nash equilibrium is a set of Bellman and policy solutions for the set of players

Discrete-Time Algorithm

Order the states in Ω

Make initial guesses for the value $V^i(\omega)$ and the policy $X^i(\omega)$ of each player $i = 1, \ldots, N$ in each state $\omega \in \Omega$.

For each state $\omega \in \Omega$, use $V^i(\omega)$ and $X^i(\omega)$ to compute new guesses $\hat{V}^i(\omega)$ and $\hat{X}^i(\omega)$ for each player $i = 1, \ldots, N$

$$
\begin{aligned}
\hat{X}^i &\leftarrow \arg\max_{x^i} \pi^i\left(x^i, X^{-i}, \omega\right) \\
&\quad + \beta \mathrm{E}_{\omega'}\left\{\Phi^i\left(x^i, X^{-i}, \omega, \omega'\right) + V^i\left(\omega'\right) | \omega, x^i, X^{-i}\right\}, \qquad (1) \\
\hat{V}^i &\leftarrow \pi^i\left(\hat{X}^i, X^{-i}, \omega\right) \\
&\quad + \beta \mathrm{E}_{\omega'}\left\{\Phi^i\left(\hat{X}^i, X^{-i}, \omega, \omega'\right) + V^i\left(\omega'\right) | \omega, \hat{X}^i, X^{-i}\right\}. \qquad (2)
\end{aligned}
$$

pre-Gauss-Jacobi at each state $\omega \in \Omega$ and each player $i = 1, \ldots, N$

Does NOT compute Nash equilibrium at state $\omega \in \Omega$, only best replies.

Computational Challenge

- Equations (1) and (2) includes (set $\Phi^i = 0$ to stay simple)

$$\mathrm{E}_{\omega'}\left\{V^i\left(\omega'\right)|\omega, X(\omega)\right\} = \sum_{\{\omega':\Pr(\omega'|\omega,X(\omega))>0\}} V^i\left(\omega'\right)\Pr\left(\omega'|\omega, X(\omega)\right), \tag{3}$$

- Number of terms is all states ω' s.t. $\Pr\left(\omega'|\omega, X(\omega)\right) > 0$.
  - Independent jumps, only {up, down, no change}, sum has 3^N terms,

$$\sum_{\left\{\omega':(\omega')^i\in\left\{\omega^i-1,\omega^i,\omega^i+1\right\},i=1,\ldots,N\right\}} V^i\left(\omega'\right)\prod_{i=1}^{N}\overset{i}{\Pr}\left(\left(\omega'\right)^i|\omega^i, X^i(\omega)\right).$$

  - Each player can move K states implies K^N terms.
  - CURSE OF DIMENSIONALITY!!!

Continuous-time Model

Same model but with continuous time.

- Each firm's path is a piecewise-constant, right-continuous function of time.
- Jumps occur at random times, following a controlled Poisson process.
  - At t hazard rate of a jump is $\phi(x_t, \omega_t)$.
  - If jump at t, prob. state moves to ω' is $f(\omega' | \omega_{t-}, x_{t-})$, where $\omega_{t-} = \lim_{s \to t-} \omega_s$, and $x_{t-} = \lim_{s \to t-} x_s$.
  - $f(\omega' | \omega_{t-}, x_{t-})$ represents the induced first-order Markov process.
  - WLOG, $f(\omega_{t-} | \omega_{t-}, x_{t-}) = 0$.

- Over a short interval of time $\Delta > 0$

$$\begin{aligned} \Pr\left(\omega_{t+\Delta} \neq \omega_t | \omega_t, x_t\right) &= \phi\left(x_t, \omega_t\right)\Delta + O\left(\Delta^2\right), \\ \Pr\left(\omega_{t+\Delta} = \omega' | \omega_t, x_t, \omega_{t+\Delta} \neq \omega_t\right) &= f\left(\omega' | \omega_t, x_t\right) + O\left(\Delta\right). \end{aligned}$$

- Independent transitions implies

$$\begin{aligned} \overset{i}{\Pr}\left(\omega^i_{t+\Delta} \neq \omega^i_t | \omega^i_t, x^i_t\right) &= \phi^i\left(x^i_t, \omega^i_t\right)\Delta + O\left(\Delta^2\right), \\ \overset{i}{\Pr}\left(\omega^i_{t+\Delta} = \left(\omega'\right)^i | \omega^i_t, x^i_t, \omega^i_{t+\Delta} \neq \omega^i_t\right) &= f^i\left(\left(\omega'\right)^i | \omega^i_t, x^i_t\right) + O\left(\Delta\right), \end{aligned}$$

  where $\phi\left(x_t, \omega_t\right) = \sum_{i=1}^N \phi^i\left(x^i_t, \omega^i_t\right)$ is hazard rate of some jump
- Key fact: during short Δ, at most one jump by one firm, a.s.

Payoff of player i consists of two components.

- payoff flow equal to $\pi^i(x_t, \omega_t)$, a flow
- $\Phi^i(x_{t-}, \omega_{t-}, \omega_t)$ is the jump in player wealth if jump

Objective of player i is

$$E\left\{\int_0^\infty e^{-\rho t}\pi^i(x_t, \omega_t)\,dt + \sum_{m=1}^{\infty} e^{-\rho T_m}\Phi^i\left(x_{T_m^-}, \omega_{T_m^-}, \omega_{T_m}\right)\right\}, \qquad (4)$$

Bellman equation for player i over a short interval of time of length $\Delta > 0$:

$$\begin{aligned} V^i &= \max_{x^i} \pi^i\left(x^i, X^{-i}, \omega\right)\Delta \\ &\quad + (1-\rho\Delta)\left\{\left(1-\phi\left(x^i, X^{-i}, \omega\right)\Delta - O\left(\Delta^2\right)\right)V^i\right. \\ &\quad + \left(\phi\left(x^i, X^{-i}, \omega\right)\Delta + O\left(\Delta^2\right)\right) \\ &\qquad \left.\times\left(E_{\omega'}\left\{\Phi^i\left(x^i, X^{-i}, \omega, \omega'\right) + V^i(\omega')\,|\omega, x^i, X^{-i}\right\} + O(\Delta)\right)\right\}, \end{aligned}$$

which, as $\Delta \to 0$, simplifies to the Bellman equation

$$\begin{aligned}\rho V^i(\omega) = \max_{x^i \in \mathbb{X}^i(\omega)} \ & \pi^i(x^i, X^{-i}(\omega), \omega) - \phi(x^i, X^{-i}(\omega), \omega) V^i(\omega) \\ & + \phi(x^i, X^{-i}(\omega), \omega) \\ & \times \mathrm{E}_{\omega'}\{\Phi^i(x^i, X^{-i}(\omega), \omega, \omega') + V^i(\omega') | \omega, x^i, X^{-i}(\omega)\}.\end{aligned} \tag{4}$$

Hence, $V^i(\omega)$ can be interpreted as the asset value to player i of participating in the game. This asset is priced by requiring that the opportunity cost of holding it, $\rho V^i(\omega)$, equals the current cash flow, $\pi^i(x^i, X^{-i}(\omega), \omega)$, plus the expected capital gain or loss conditional on a jump occurring,

$$\mathrm{E}_{\omega'}\{\Phi^i(x^i, X^{-i}(\omega), \omega, \omega') + V^i(\omega') | \omega, x^i, X^{-i}(\omega)\} - V^i(\omega),$$

3.2 *Continuous-time algorithm*

In its basic form, our computational strategy adapts the block Gauss–Seidel scheme to the continuous-time model. The sole change is that to update players' values and policies in state $\omega \in \Omega$, we replace equations (9) and (10) by

$$\begin{aligned} \hat{X}^i(\omega) \leftarrow \arg\max_{x^i} \pi^i(x^i, X^{-i}(\omega), \omega) - \phi(x^i, X^{-i}(\omega), \omega) V^i(x^i, X^{-i}(\omega), \omega) \\ + \phi(x^i, X^{-i}(\omega), \omega) \\ \times \mathrm{E}_{\omega'}\left\{\Phi^i(x^i, X^{-i}(\omega), \omega, \omega') + V^i(\omega') | \omega, x^i, X^{-i}(\omega)\right\}, \end{aligned} \tag{11}$$

$$\begin{aligned} \hat{V}^i(\omega) \leftarrow \frac{1}{\rho + \phi(\hat{X}^i(\omega), X^{-i}(\omega), \omega)} \pi^i(\hat{X}^i(\omega), X^{-i}(\omega), \omega) \\ + \frac{\phi(\hat{X}^i(\omega), X^{-i}(\omega), \omega)}{\rho + \phi(\hat{X}^i(\omega), X^{-i}(\omega), \omega)} \\ \times \mathrm{E}_{\omega'}\left\{\Phi^i(\hat{X}^i(\omega), X^{-i}(\omega), \omega, \omega') + V^i(\omega') | \omega, \hat{X}^i(\omega), X^{-i}(\omega)\right\}. \end{aligned} \tag{12}$$

Computational considerations

- Equilibrium is a finite set of equations, each equation being a low-dimensional optimization problem
- LOOKS like dynamic programming but it is not
  - This is not a contraction mapping
  - There may be multiple solutions, in which case this cannot be a contraction mapping
  - Without a contraction factor you cannot use simple stopping rule form DP
- The system is a set of nonlinear equations
  - Can use Gauss-Jacobi, as did Pakes and Mcguire
  - Could use Gauss-Seidel, as later people did (to save memory)
  - Different algorithms may produce different solutions

TABLE 1. Time per iteration per state per firm and percentage of time spent on computing the expectation.[a]

Number of Firms	Number of States	Number of Unknowns	Discrete Time		Continuous Time Without Precomputed Addresses		Continuous Time With Precomputed Addresses		Ratio: Discrete to Continuous Time Without Precomputed Addresses	Ratio: Continuous Time Without to With Precomputed Addresses	Ratio: Discrete to Continuous Time With Precomputed Addresses
			sec	%	sec	%	sec	%			
2	171	684	1.07 (−6)	55	7.13 (−7)	41	5.85 (−7)	36	1.50	1.22	1.83
3	1140	6840	1.61 (−6)	76	6.67 (−7)	44	5.26 (−7)	38	2.41	1.27	3.06
4	5985	47,880	3.30 (−6)	87	6.68 (−7)	49	5.10 (−7)	41	4.94	1.31	6.48
5	26,334	263,340	8.05 (−6)	98	7.06 (−7)	49	5.24 (−7)	43	11.40	1.35	15.36
6	100,947	1,211,364	2.15 (−5)	97	7.51 (−7)	52	5.37 (−7)	46	28.57	1.40	40.00
7	346,104	4,845,456	6.19 (−5)	100	7.74 (−7)	56	5.47 (−7)	49	80.00	1.42	113.21
8	1,081,575	17,305,200	1.65 (−4)	100	8.23 (−7)	58	5.92 (−7)	56	200.28	1.39	278.44

[a] Quality ladder model with $M = 18$ quality levels per firm and a discount factor of 0.925. (k) is shorthand for $\times 10^k$.

TABLE 2. Time per iteration per state per firm and percentage of time spent on computing the expectation.[a]

Number of Firms	Number of States	Number of Unknowns	Discrete Time		Continuous Time Without Precomputed Addresses		Continuous Time With Precomputed Addresses		Ratio		
									Disc. to Cont. Time Without Precomp. Addresses	Cont. Time Without to With Precomp. Addresses	Disc. to Cont. Time With Precomp. Addresses
			sec	%	sec	%	sec	%			
2	45	180	9.78 (−7)	52	6.89 (−7)	42	5.67 (−7)	33	1.42	1.22	1.73
3	165	990	1.45 (−6)	74	6.36 (−7)	44	5.05 (−7)	38	2.29	1.26	2.88
4	495	3960	2.90 (−6)	88	6.36 (−7)	48	4.75 (−7)	43	4.55	1.34	6.10
5	1287	12,870	6.94 (−6)	96	6.42 (−7)	53	4.77 (−7)	46	10.81	1.35	14.57
6	3003	36,036	1.81 (−5)	98	6.88 (−7)	55	4.88 (−7)	45	26.34	1.41	37.12
7	6435	90,090	5.02 (−5)	100	7.33 (−7)	53	5.11 (−7)	48	68.48	1.43	98.26
8	12,870	205,920	1.31 (−4)	100	7.77 (−7)	55	5.24 (−7)	50	168.33	1.48	249.38
9	24,310	437,580	3.82 (−4)	100	7.77 (−7)	62	5.39 (−7)	53	492.16	1.44	709.04
10	43,758	875,160	1.07 (−3)	100	8.34 (−7)	64	5.94 (−7)	44	1282.19	1.40	1800.00
11	75,582	1,662,804	2.99 (−3)	100	8.42 (−7)	67	5.77 (−7)	56	3557.14	1.46	5187.50
12	125,970	3,023,280	8.20 (−3)	100	8.60 (−7)	68	5.95 (−7)	60	9533.08	1.44	13,770.00
13	203,490	5,290,740	2.42 (−2)	100	9.22 (−7)	69	6.20 (−7)	61	26,235.65	1.49	39,033.56
14	319,770	8,953,560	6.76 (−2)	100	9.53 (−7)	72	6.55 (−7)	59	70,946.70	1.45	103,195.27

[a]Quality ladder model with $M = 9$ quality levels per firm and a discount factor of 0.925. (k) is shorthand for $\times 10^k$.

Table 3. Number of iterations to convergence.[a]

Number of Firms	Discount Factor	Discrete Time		Continuous Time		Ratio	
		$<10^{-4}$	$<10^{-8}$	$<10^{-4}$	$<10^{-8}$	$<10^{-4}$	$<10^{-8}$
3	0.925	118	201	212	446	0.56	0.45
3	0.98	412	702	776	1699	0.53	0.41
3	0.99	782	1367	1531	3393	0.51	0.40
3	0.995	1543	2719	3042	6779	0.51	0.40
6	0.925	118	201	364	725	0.32	0.28
6	0.98	494	780	1674	3324	0.30	0.23
6	0.99	983	1525	3379	6761	0.29	0.23
6	0.995	1900	2945	6797	13,637	0.28	0.22
9	0.925	119	201	404	818	0.29	0.25
9	0.98	492	775	2363	4493	0.21	0.17
9	0.99	988	1526	4973	9469	0.20	0.16
9	0.995	2003	3042	10,148	19,365	0.20	0.16
12	0.925			412	854		
12	0.98			2721	5106		
12	0.99			6023	11,181		
12	0.995			12,580	23,304		

[a]The stopping rule is either "distance to truth $<10^{-4}$" or "distance to truth $<10^{-8}$." Quality ladder model with $M = 9$ quality levels per firm.

TABLE 4. Time to convergence.[a]

Number of Firms	Discrete Time (min)	Continuous Time (min)	Ratio		
			Time per Iteration	Number of Iterations	Time to Convergence
2	1.80 (−4)	1.12 (−4)	1.73	0.93	1.61
3	1.42 (−3)	8.83 (−4)	2.88	0.56	1.60
4	1.13 (−2)	4.43 (−3)	6.10	0.42	2.54
5	8.78 (−2)	1.70 (−2)	14.57	0.36	5.18
6	6.42 (−1)	5.34 (−2)	37.12	0.32	12.03
7	4.44 (0)	1.47 (−1)	98.26	0.31	30.19
8	2.67 (1)	3.56 (−1)	249.38	0.30	74.94
9	1.66 (2)	7.95 (−1)	709.04	0.29	208.85
10	*9.28 (2)*	1.77 (0)	1800.00	*0.29*	*523.72*
11	*4.94 (3)*	3.30 (0)	5187.50	*0.29*	*1498.33*
12	*2.46 (4)*	6.18 (0)	13,770.00	*0.29*	*3977.26*
13	*1.27 (5)*	1.13 (1)	39,033.56	*0.29*	*11,246.96*
14	*6.00 (5)*	2.02 (1)	103,195.27	*0.29*	*29,734.23*

[a]The stopping rule is “distance to truth $<10^{-4}$.” Entries in italics are based on an estimated 119 iterations to convergence in discrete time. Quality ladder model with $M = 9$ quality levels per firm and a discount factor of 0.925. (k) is shorthand for $\times 10^k$.

TABLE 5. Stopping rules.[a]

		Ad Hoc Rule		Adaptive Rule		
Number of Firms	Discount Factor	Terminal Iteration	Distance to Truth	Terminal Iteration	Distance to Truth	Convergence Factor
3	0.925	131	2.43 (−3)	218	7.84 (−5)	0.9676
3	0.98	313	9.93 (−3)	775	1.00 (−4)	0.9899
3	0.99	455	2.01 (−2)	1483	1.27 (−4)	0.9937
3	0.995	589	4.05 (−2)	2778	1.92 (−4)	0.9953
6	0.925	220	3.84 (−3)	370	8.42 (−5)	0.9777
6	0.98	742	1.78 (−2)	1689	9.17 (−5)	0.9948
6	0.99	1198	3.66 (−2)	3454	8.15 (−5)	0.9978
6	0.995	1832	7.41 (−2)	6766	1.04 (−4)	0.9986
9	0.925	232	4.45 (−3)	407	9.18 (−5)	0.9791
9	0.98	1100	2.30 (−2)	2387	9.01 (−5)	0.9961
9	0.99	1927	4.87 (−2)	5091	7.84 (−5)	0.9984
9	0.995	3129	1.00 (−1)	10,358	8.10 (−5)	0.9992
12	0.925	227	4.73 (−3)	411	1.02 (−4)	0.9781
12	0.98	1276	2.58 (−2)	2751	8.89 (−5)	0.9966
12	0.99	2447	5.59 (−2)	6185	7.48 (−5)	0.9987
12	0.995	4217	1.16 (−1)	12,994	7.01 (−5)	0.9994

[a]Terminal iteration, distance to truth at terminal iteration, and estimated convergence factor are shown. Prespecified tolerance is 10^{-4}. Continuous-time quality ladder model with $M = 9$ quality levels per firm. (k) is shorthand for $\times 10^{k}$.

Computational Gains

- Curse of dimensionality in number of firms with discrete time
- NO curse with continuous time

More Computational considerations

- Parallelization?
  - Much more dangerous, but should be tried
  - Gauss-Jacobi is likely less dangerous than Gauss-Seidel