Dynamic Programming

Kenneth L. Judd

(Joint work with Yongyang Cai)

September 6, 2011

Dynamic Programming

Powerful for solving dynamic stochastic optimization problems

- Based on principle of recursion due to Bellman and Isaacs
- Replaces multiperiod optimization problems with a sequence of two-period problems

Applications

- Economics
  - Business investment
  - Life-cycle decisions on labor, consumption, education, portfolio choice
  - Economic policy
- Operations Research
  - Scheduling, queueing
  - Inventory
- Climate change
  - Economic response to climate policies
  - Optimal policy response to global warming problems

Canonical Example in Economics

General Stochastic Accumulation

- Problem:

$$
\begin{aligned}
V(k,\theta) = \max_{c_t,\ell_t} \; E\left\{\sum_{t=0}^{\infty} \beta^t \; u(c_t,\ell_t)\right\} \\
k_{t+1} = F(k_t,\ell_t,\theta_t) - c_t \\
\theta_{t+1} = g(\theta_t,\varepsilon_t) \\
k_0 = k,\; \theta_0 = \theta.
\end{aligned}
$$

- State variables:
  - k: productive capital stock, endogenous
  - θ: productivity state, exogenous
- The dynamic programming formulation is

$$
V(k,\theta) = \max_{c,\ell} \; u(c,\ell) + \beta E\{V(F(k,\ell,\theta) - c, \theta^{+})|\theta\}, \quad (12.1.21)
$$

where θ^{+} is next period's θ realization.

Definitions

Discrete-Time Dynamic Programming

- Objective:

$$E\left\{\sum_{t=1}^{T}\pi(x_t,\ u_t,\ t)+W(x_{T+1})\right\} \tag{1}$$

  - X: set of states
  - $\mathcal{D}$: the set of controls
  - $\pi(x,\ u,\ t)$ payoffs in period t, for $x\in X$ at the beginning of period t, and control $u\in\mathcal{D}$ is applied in period t.
  - $D(x,t)\subseteq\mathcal{D}$: controls which are feasible in state x at time t.
  - $F(A;x,u,t)$: probability that $x_{t+1}\in A\subset X$ conditional on time t control and state
- Value function

$$V(x,t)\equiv\sup_{\mathcal{U}(x,t)}E\left\{\sum_{s=t}^{T}\pi(x_s,\ u_s,\ s)+W(x_{T+1})|x_t=x\right\} \tag{2}$$

- Bellman equation

$$V(x,t)=\sup_{u\in D(x,t)}\ \pi(x,\ u,\ t)+E\left\{V(x_{t+1},\ t+1)|x_t=x,u_t=u\right\} \tag{3}$$

- Existence: boundedness of π is sufficient

Parametric Approach

General Parametric Approach: Approximating T

- For each x_j, $(TV)(x_j)$ is defined by

$$v_j = (TV)(x_j) = \max_{u \in D(x_j)} \pi(u, x_j) + \beta \int \hat{V}(x^+; a) dF(x^+ | x_j, u) \quad (4)$$

- In practice, we compute the approximation $\hat{T}$

$$v_j = (\hat{T}V)(x_j) \doteq (TV)(x_j)$$

  - Integration step: for ω_j and x_j for some numerical quadrature formula

$$\begin{aligned} E\{V(x^+; a) | x_j, u)\} &= \int \hat{V}(x^+; a) dF(x^+ | x_j, u) \\ &= \int \hat{V}(g(x_j, u, \varepsilon); a) dF(\varepsilon) \\ &\doteq \sum_\ell \omega_\ell \hat{V}(g(x_j, u, \varepsilon_\ell); a) \end{aligned}$$

  - Maximization step: for $x_i \in X$, evaluate

$$v_i = (T\hat{V})(x_i)$$

  - Fitting step:
    - Data: (v_i, x_i), $i = 1, \cdots, n$
    - Objective: find an $a \in R^m$ such that $\hat{V}(x; a)$ best fits the data

Shape-preserving Chebyshev Interpolation

Problem: Instability of Value Function Iteration
Solution: LP model for shape-preserving Chebyshev Interpolation:

$$\begin{aligned}
\min_{c_j} \quad & \sum_{j=0}^{m-1} (c_j^+ + c_j^-) + \sum_{j=m}^{n} (j+1-m)^2 (c_j^+ + c_j^-) \\
\text{s.t.} \quad & \sum_{j=0}^{n} c_j T_j'(y_i) > 0 > \sum_{j=0}^{n} c_j T_j''(y_i), \qquad i = 1, \ldots, m', \\
& \sum_{j=0}^{n} c_j T_j(z_i) = v_i, \quad i = 1, \ldots, m, \\
& c_j - \hat{c}_j = c_j^+ - c_j^-, \quad j = 0, \ldots, m-1, \\
& c_j = c_j^+ - c_j^-, \quad j = m, \ldots, n, \\
& c_j^+ \geq 0, \quad c_j^- \geq 0, \quad j = 1, \ldots, n.
\end{aligned}$$

Optimal Growth Example

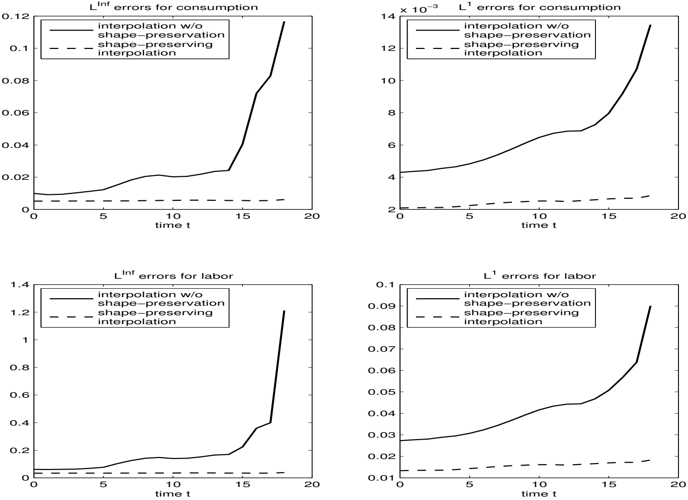
- Optimal Growth Problem:

$$V_0(k_0) = \max_{c,l} \sum_{t=0}^{T-1} \beta^t u(c_t, l_t) + \beta^T V_T(k_T),$$
$$\text{s.t.} \quad k_{t+1} = F(k_t, l_t) - c_t, \quad 0 \le t < T$$

- DP model of optimal growth problem:

$$V_t(k) = \max_{c,l} \ u(c, l) + \beta V_{t+1}(F(k, l) - c)$$

Errors of NDP with Chebyshev interpolation (shape-preserving or not)



Portfolio Optimization Example

- W_t: wealth at stage t; stocks' random return: $R = (R_1, \ldots, R_n)$; bond's riskfree return: R_f;
- $S_t = (S_{t1}, \ldots, S_{tn})^\top$: money in the stocks; $B_t = W_t - e^\top S_t$: money in the bond,
- $W_{t+1} = R_f(W_t - e^\top S_t) + R^\top S_t$
- Multi-Stage Portfolio Optimization Problem:

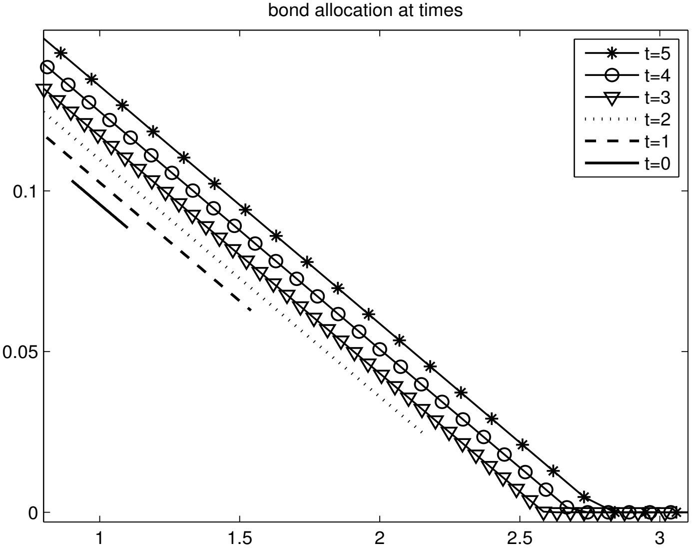
$$V_0(W_0) = \max_{X_{\boldsymbol{t}}, 0 \le t < T} E\{u(W_T)\}$$

- Bellman Equation:

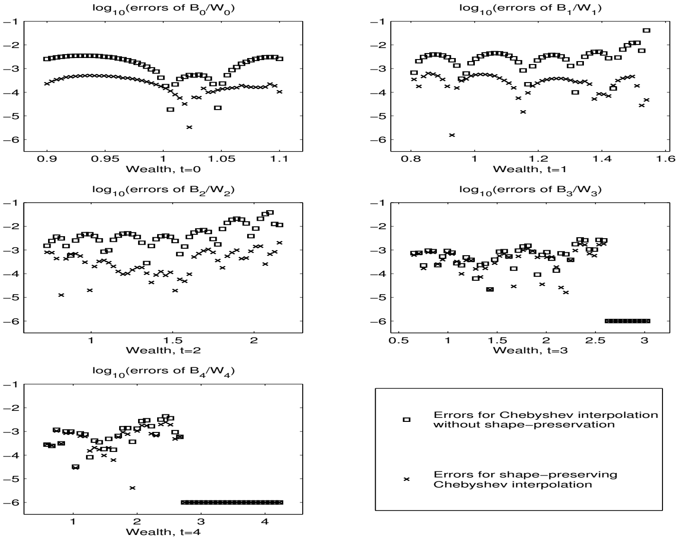
$$V_t(W) = \max_S \ E\{V_{t+1}(R_f(W - e^\top S) + R^\top S)\}$$

W: state variable; S: control variables.

Exact optimal bond allocation



Errors of Optimal Stock Allocations (shape-preserving or not)



Derivative of Value Functions in General Models

- For an optimization problem,

$$\begin{aligned} V(x) &= \max_{y} \; f(x,y) \\ & \quad \text{s.t. } g(x,y)=0, h(x,y) \geq 0, \end{aligned}$$

  add a trivial control variable z and a trivial constraint $x-z=0$:

$$\begin{aligned} V(x) &= \max_{y,z} \; f(z,y) \\ & \quad \text{s.t. } g(z,y)=0, h(z,y) \geq 0, x-z=0. \end{aligned}$$

- Then by the envelope theorem, we get

$$V'(x) = \lambda,$$

  where λ is the shadow price for the trivial constraint $x-z=0$.
- Idea: use shadow price as new information in approximating value functions

			error of c_0^*		error of l_0^*	
γ	η	m	Lagrange	Hermite	Lagrange	Hermite
0.5	0.1	5	1.1(−1)	1.2(−2)	1.9(−1)	1.8(−2)
		10	6.8(−3)	3.1(−5)	9.9(−3)	4.4(−5)
		20	2.3(−5)	1.5(−6)	3.2(−5)	2.3(−6)
0.5	1	5	1.4(−1)	1.4(−2)	6.1(−2)	5.6(−3)
		10	7.7(−3)	3.7(−5)	3.1(−3)	1.6(−5)
		20	2.6(−5)	6.5(−6)	1.1(−5)	3.0(−6)
2	0.1	5	5.5(−2)	6.1(−3)	2.7(−1)	3.6(−2)
		10	3.5(−3)	2.1(−5)	2.0(−2)	1.2(−4)
		20	1.6(−5)	1.4(−6)	9.1(−5)	7.6(−6)
2	1	5	9.4(−2)	1.1(−2)	1.3(−1)	1.7(−2)
		10	5.7(−3)	3.9(−5)	9.2(−3)	6.1(−5)
		20	2.8(−5)	4.7(−6)	4.3(−5)	8.0(−6)
8	0.1	5	2.0(−2)	2.2(−3)	3.6(−1)	4.9(−2)
		10	1.2(−3)	8.5(−6)	2.7(−2)	1.9(−4)
		20	6.1(−6)	1.0(−6)	1.4(−4)	4.4(−6)
8	1	5	6.6(−2)	7.2(−3)	3.4(−1)	4.5(−2)
		10	3.0(−3)	2.6(−5)	2.0(−2)	1.7(−4)
		20	2.0(−5)	0.0(−7)	1.3(−4)	2.1(−7)

Note: $a(k)$ means $a \times 10^k$.

Shape-preserving Hermite Spline Interpolation

- Idea: impose shape and use gradient information
- Using Hermite data $\{(x_i, v_i, s_i) : i = 1, \ldots, m\}$,

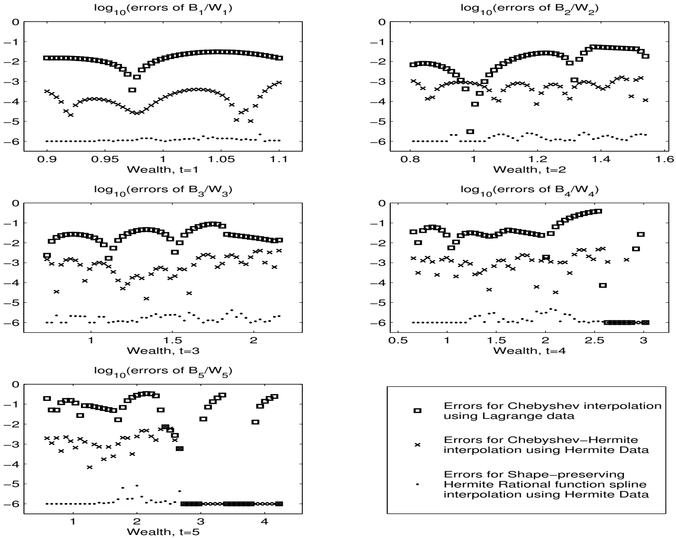
$$\hat{V}(x; \mathbf{c}) = c_{i1} + c_{i2}(x - x_i) + \frac{c_{i3}c_{i4}(x - x_i)(x - x_{i+1})}{c_{i3}(x - x_i) + c_{i4}(x - x_{i+1})},$$

when $x \in [x_i, x_{i+1}]$, where

$$\begin{aligned} c_{i1} &= v_i, \\ c_{i2} &= \frac{v_{i+1} - v_i}{x_{i+1} - x_i}, \\ c_{i3} &= s_i - c_{i2}, \\ c_{i4} &= s_{i+1} - c_{i2}, \end{aligned}$$

for $i = 1, \ldots, m - 1$.

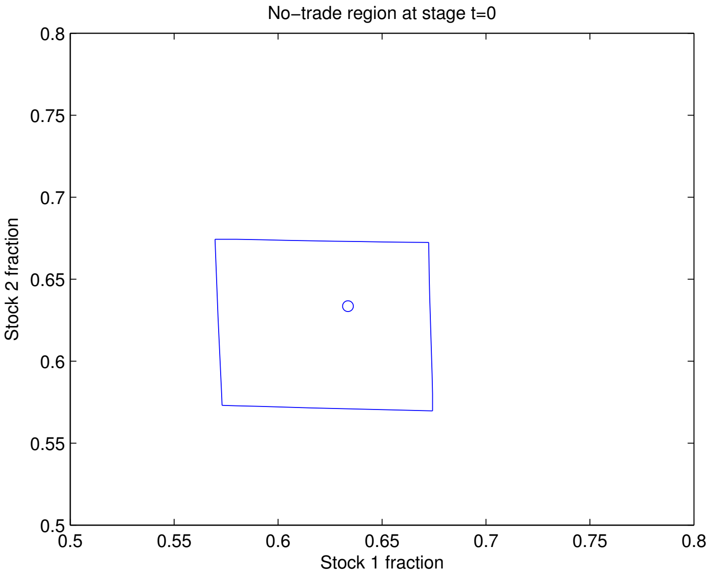
Errors of Optimal Bond Allocations (Lagrange vs Hermite vs Shape-preserving+Hermite)



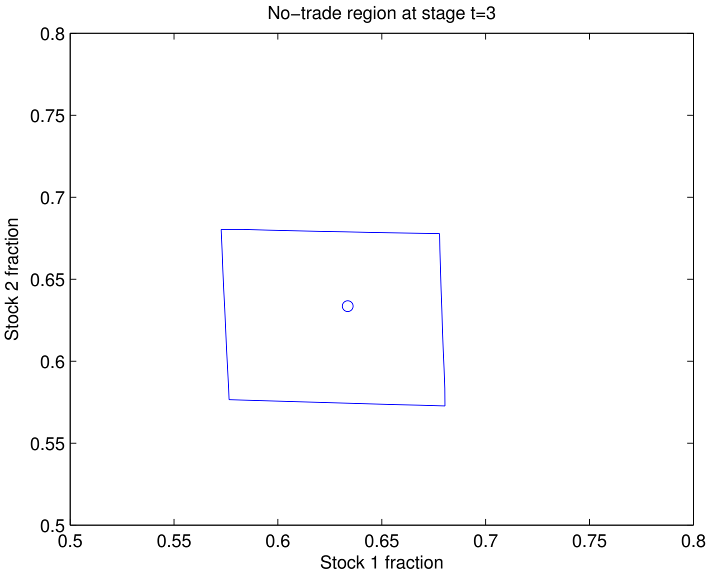
Proportional Transaction Cost and CRRA Utility

- Separability of wealth W and portfolio fractions x.
- If $u(W) = W^{1-\gamma}/(1-\gamma)$, then $V_t(W_t, x_t) = W_t^{1-\gamma} \cdot g_t(x_t)$.
- If $u(W) = \log(W)$, then $V_t(W_t, x_t) = \log(W_t) + \psi_t(x_t)$.
- "No-trade" region: $\Omega_t = \{x_t : \ (\delta_t^+)^* = (\delta_t^-)^* = 0\}$, where $(\delta_t^+)^* \geq 0$ are fractions of wealth for buying stocks, and $(\delta_t^-)^* \geq 0$ are fractions of wealth for selling stocks.

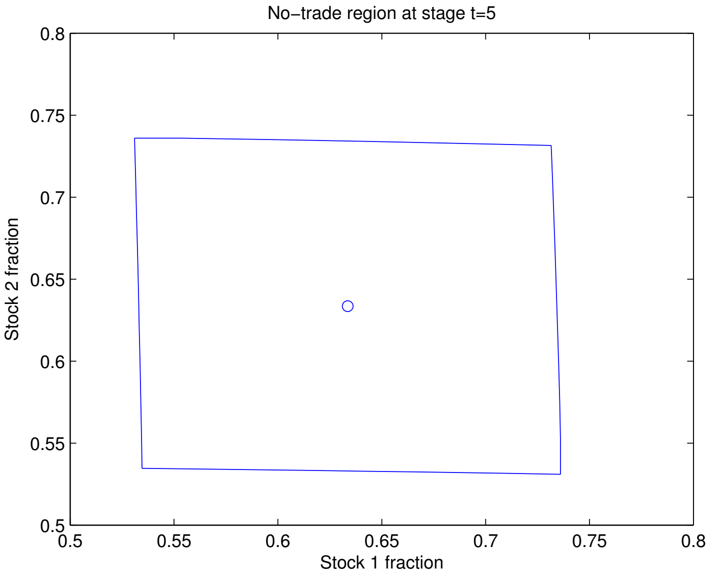
2 stocks with i.i.d. returns at $t = 0$ (liquidate at $t = 6$)



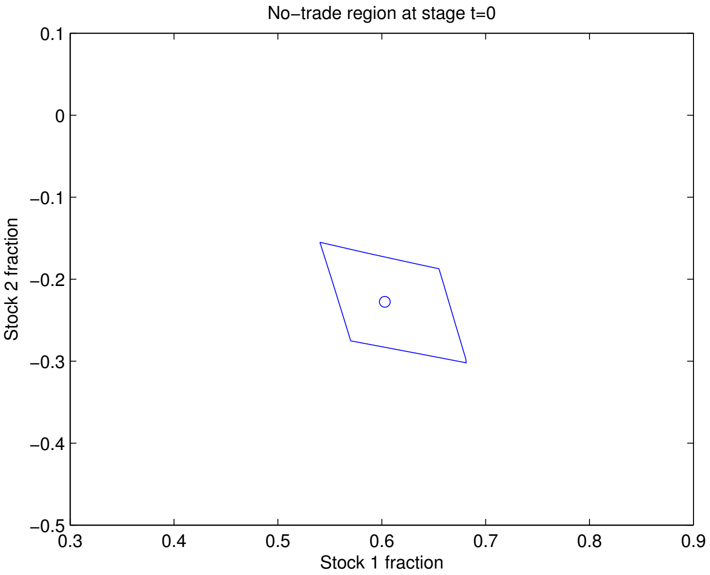
2 stocks with i.i.d. returns at $t = 3$ (liquidate at $t = 6$)



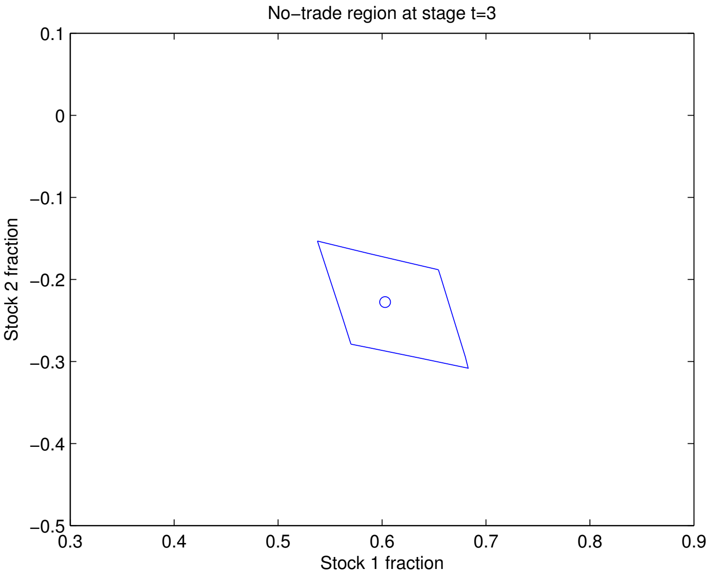
2 stocks with i.i.d. returns at $t = 5$ (liquidate at $t = 6$)



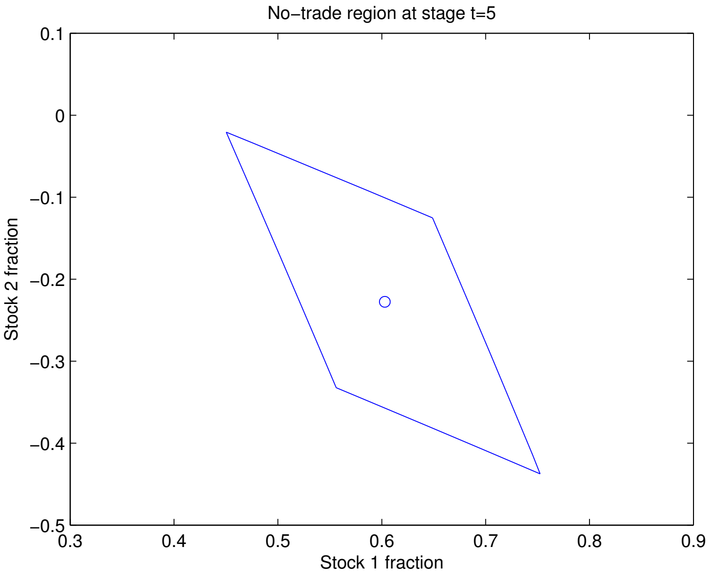
2 stocks with correlated returns at $t = 0$ (liquidate at $t = 6$)



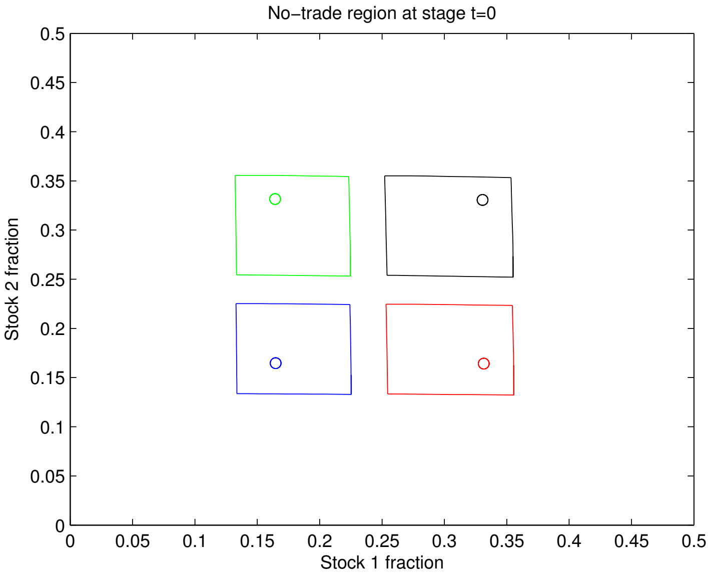
2 stocks with correlated returns at $t = 3$ (liquidate at $t = 6$)



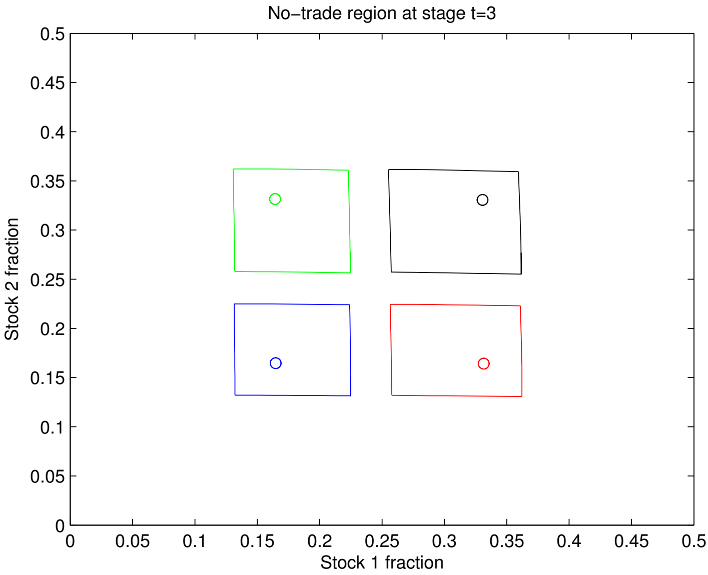
2 stocks with correlated returns at $t = 5$ (liquidate at $t = 6$)



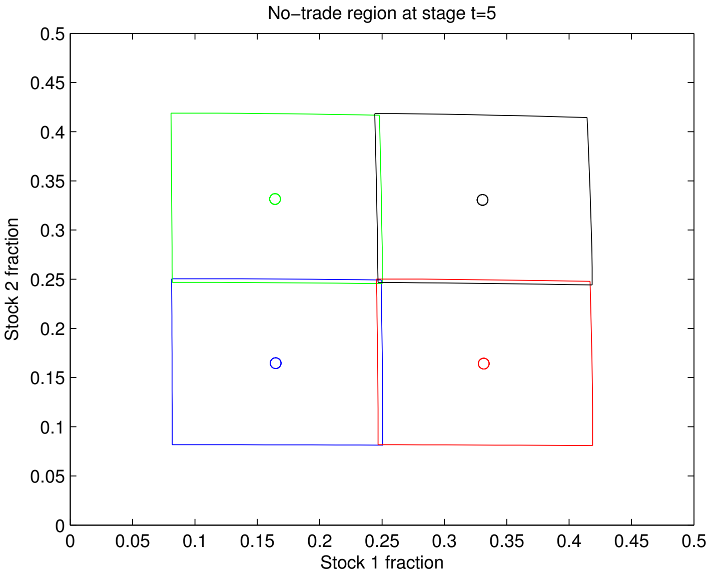
2 stocks with stochastic μ at $t = 0$ (liquidate at $t = 6$)



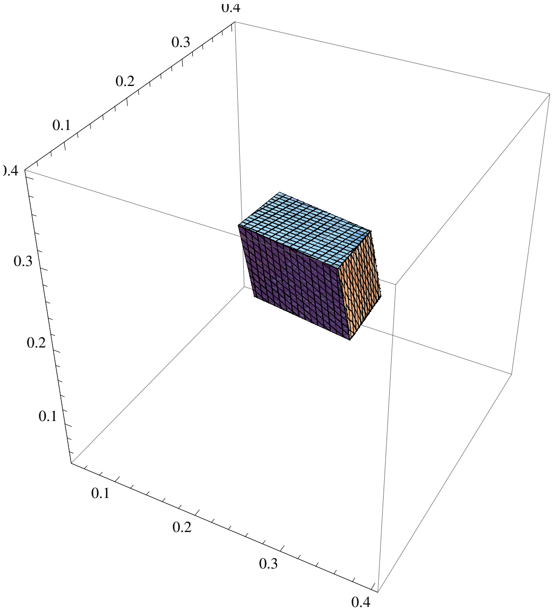
2 stocks with stochastic μ at $t = 3$ (liquidate at $t = 6$)



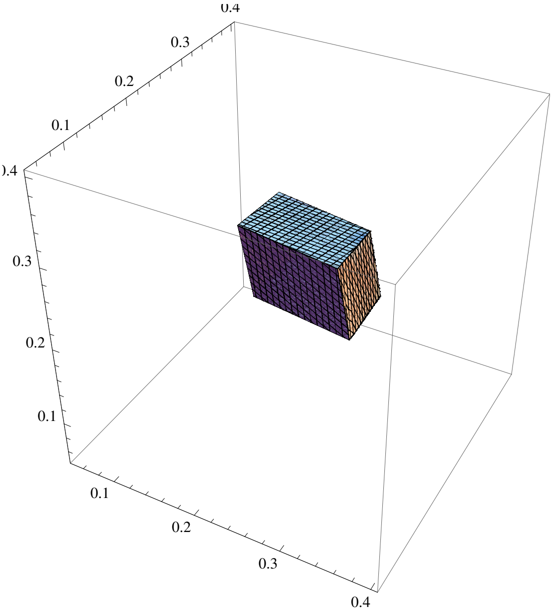
2 stocks with stochastic μ at $t = 5$ (liquidate at $t = 6$)



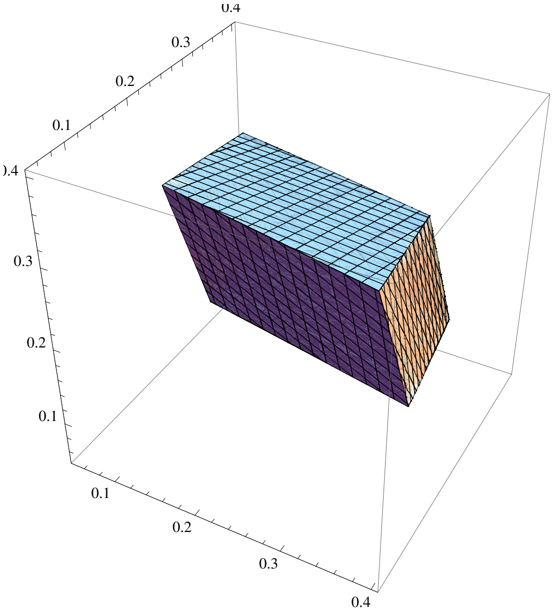
3 correlated stocks at $t = 0$ (liquidate at $t = 6$)



3 correlated stocks at $t = 3$ (liquidate at $t = 6$)



3 correlated stocks at $t = 5$ (liquidate at $t = 6$)

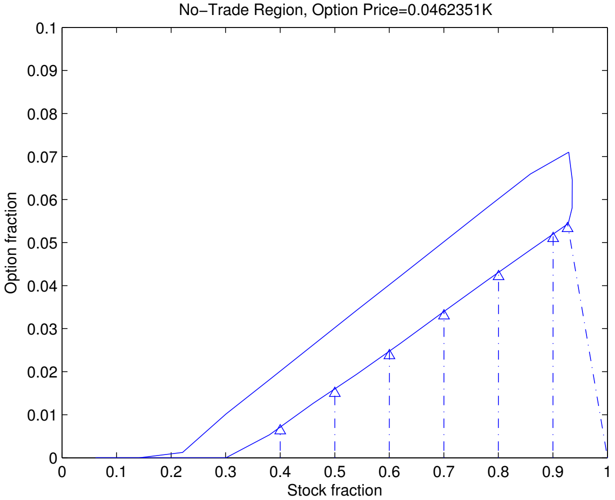


Application of Portfolio Analysis

Options

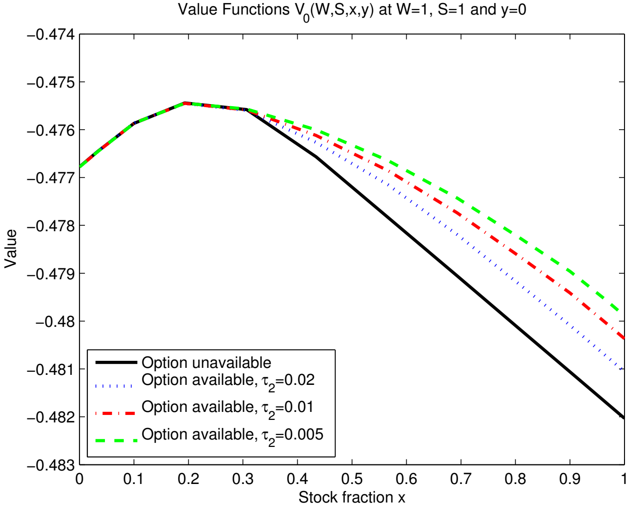
- The pricing theory of options assumes that options have no social value
- Finance people claim that they economize on transaction costs, but provide no analysis
- Cai has shown that there is some value to one option; future work will examine social value of free entry

1 stock and 1 at-the-money put option at $t = 0$ (liquidate at $t = 6$ months)



- **Put option: strike** K**, expiration time** T**, payoff** $\max(K - S_T, 0)$
- **stock price** S**, utility** $u(W) = -W^{-2}/2$

Value functions with/without options at $t = 0$ (liquidate at $t = 6$)



- (x, y): **fractions of money in stock and option**
- $\tau_1 = 0.01$ **and** τ_2: **transaction cost ratios of stock and option**

Introduction

- All IAMs (Integrated Assessment Models) are deterministic
- Most are myopic, not forward-looking
- This combination makes it impossible for IAMs to consider decisions in a dynamic, evolving and uncertain world
- We formulate dynamic stochastic general equilibrium extensions of DICE (Nordhaus)
- Conventional wisdom: *"Integration of DSGE models with long run intertemporal models like IGEM is beyond the scientific frontier at the moment"* (Peer Review of ADAGE and IGEM, June 2010)
- Fact: We use multidimensional dynamic programming methods, developed over the past 20 years in Economics, to study dynamically optimal policy responses

DSICE

Cai-Judd-Lontzek DSICE Model:
Dynamic Stochastic Integrated Model of Climate and Economy

$$\begin{aligned} DSICE &= DICE2007 \\ &- \text{constraint on savings rate }, i.e.: s = .22 \\ &- \text{ad hoc finite difference method} \\ &+ \text{stochastic production function} \\ &+ \text{stochastic damage function} \\ &+ \text{1-year period length} \end{aligned}$$

stochastic means: intrinsic random events within the specific model, not uncertain parameters

- DSICE: solve stochastic optimization problem

$$\max_{c_t, l_t, \mu_t} \mathbb{E}\left\{\sum_{t=0}^{\infty} \beta^t u(c_t, l_t)\right\}$$

$$\begin{aligned}
\text{s.t.} \quad k_{t+1} &= (1-\delta)k_t + \Omega_t(1-\Lambda_t)Y_t - c_t, \\
M_{t+1} &= \Phi^M M_t + (E_t, 0, 0)^\top, \\
T_{t+1} &= \Phi^T T_t + (\xi_1 F_t, 0)^\top, \\
\zeta_{t+1} &= g^\zeta(\zeta_t, \omega_t^\zeta), \\
J_{t+1} &= g^J(J_t, \omega_t^J)
\end{aligned}$$

- output: $Y_t \equiv f(k_t, l_t, \zeta_t, t) = \zeta_t A_t k_t^\alpha l_t^{1-\alpha}$
- damages: $\Omega_t \equiv \frac{J_t}{1+\pi_1 T_t^{AT} + \pi_2 (T_t^{AT})^2}$
- ζ_t: productivity shock, J_t: damage function shock

- emission control effort: $\Lambda_t \equiv \psi_t^{1-\theta_2}\theta_{1,t}\mu_t^{\theta_2}$
- Mass of carbon concentration: $M_t = (M_t^{AT}, M_t^{LO}, M_t^{UP})^\top$
- Temperature: $T_t = (T_t^{AT}, T_t^{LO})^\top$
- Total carbon emission: $E_t = E_{Ind,t} + E_{Land,t}$, where

$$E_{Ind,t} = \sigma_t(1-\mu_t)(f_1(k_t, l_t, \theta_t, t))$$

- Total radiative forcing (watts per square meter from 1900):

$$F_t = \eta \log_2(M_t^{AT}/M_0^{AT}) + F_t^{EX}$$

- DP model for DSICE :

$$
\begin{aligned}
V_t(k,\zeta,J,M,T) &= \max_{c,l,\mu} \quad u(c,l)+\beta\mathbb{E}[V_{t+1}(k^+,\zeta^+,J^+,M^+,T^+)] \\
\text{s.t.} \quad k^+ &= (1-\delta)k+\Omega_t(1-\Lambda_t)f(k,l,\zeta,t)-c, \\
M^+ &= \Phi^M M+(E_t,0,0)^\top, \\
T^+ &= \Phi^T T+(\xi_1 F_t,0)^\top, \\
\zeta^+ &= g^\zeta(\zeta,\omega^\zeta), \\
J^+ &= g^J(J,\omega^J)
\end{aligned}
$$

Parallel DP Algorithm

- Parallelization in Maximization step in NDP: Compute

$$v_i = \max_{a_i \in \mathcal{D}(x_i,t)} u_t(x_i, a_i) + \beta E\{\hat{V}(x_i^+; \mathbf{b}^{t+1}) | x_i, a_i\},$$

  for each $x_i \in X_t$, $1 \leq i \leq m_t$.
- Condor Master-Worker system: distributed parallelization, two entities: Master processor, a cluster of Worker processors.

Parallelization in Optimal Growth Problems

- Problem size: 4D continuous state k, 4D discrete state θ with $6^4 = 1296$ values
- Performance:

Wall clock time for all 3 VFIs	65 hours
Total time workers were up (alive)	1487 hours
Total cpu time used by all workers	1358 hours
Minimum task cpu time	557 seconds
Maximum task cpu time	4,196 seconds
Number of (different) workers	25
Overall Parallel Performance	93.56%

Parallelization in Optimal Growth Problems

Parallel efficiency for various number of worker processors

# Worker processors	Parallel efficiency	Average task CPU time (minute)	Total wall clock time (hour)
25	93.56%	**21**	**65**
54	93.46%	**25**	**33**
100	86.73%	**25**	**19**

Parallelization in Dynamic Portfolio Problems

lProblem size: 6 stocks plus 1 bond, transaction cost, number of task = 3125.

- Performance:

Wall clock time for all 6 VFIs	1.56 hours
Total time workers were up (alive)	295 hours
Total cpu time used by all workers	248 hours
Minimum task cpu time	2 seconds
Maximum task cpu time	395 seconds
Number of (different) workers	200
Overall Parallel Performance	87.2%

DSICE - DP in an Integrated Model of Climate and Economy

- All IAMs (Integrated Assessment Models) are deterministic
- Most are myopic, not forward-looking
- This combination makes it impossible for IAMs to consider decisions in a dynamic, evolving and uncertain world
- We formulate dynamic stochastic general equilibrium extensions of DICE (Nordhaus)
- Conventional wisdom: *"Integration of DSGE models with long run intertemporal models like IGEM is beyond the scientific frontier at the moment"* (Peer Review of ADAGE and IGEM, June 2010)
- Fact: We use multidimensional dynamic programming methods, developed over the past 20 years in Economics, to study dynamically optimal policy responses

DSICE

Cai-Judd-Lontzek DSICE Model:
Dynamic Stochastic Integrated Model of Climate and Economy

$$\begin{aligned} DSICE &= DICE2007 \\ &- \text{time travel for CO2} \\ &+ \text{stochastic production function} \\ &+ \text{stochastic damage function} \\ &+ \text{1-year period length} \end{aligned}$$

stochastic means: intrinsic random events within the specific model, not uncertain parameters

- DSICE: solve stochastic optimization problem

$$\max_{c_t, l_t, \mu_t} \quad \mathbb{E}\left\{\sum_{t=0}^{\infty} \beta^t u(c_t, l_t)\right\}$$

$$\begin{aligned}
\text{s.t.} \quad k_{t+1} &= (1-\delta)k_t + \Omega_t(1-\Lambda_t)Y_t - c_t, \\
M_{t+1} &= \Phi^M M_t + (E_t, 0, 0)^\top, \\
T_{t+1} &= \Phi^T T_t + (\xi_1 F_t, 0)^\top, \\
\zeta_{t+1} &= g^\zeta(\zeta_t, \omega_t^\zeta), \\
J_{t+1} &= g^J(J_t, \omega_t^J)
\end{aligned}$$

- output: $Y_t \equiv f(k_t, l_t, \zeta_t, t) = \zeta_t A_t k_t^\alpha l_t^{1-\alpha}$
- damages: $\Omega_t \equiv \frac{J_t}{1+\pi_1 T_t^{AT} + \pi_2 (T_t^{AT})^2}$
- ζ_t: productivity shock, J_t: damage function shock

- emission control effort: $\Lambda_t \equiv \psi_t^{1-\theta_2} \theta_{1,t} \mu_t^{\theta_2}$
- Mass of carbon concentration: $M_t = (M_t^{AT}, M_t^{LO}, M_t^{UP})^\top$
- Temperature: $T_t = (T_t^{AT}, T_t^{LO})^\top$
- Total carbon emission: $E_t = E_{Ind,t} + E_{Land,t}$, where

$$E_{Ind,t} = \sigma_t (1 - \mu_t)(f_1(k_t, l_t, \theta_t, t))$$

- Total radiative forcing (watts per square meter from 1900):

$$F_t = \eta \log_2(M_t^{AT} / M_0^{AT}) + F_t^{EX}$$

- DP model for DSICE :

$$\begin{aligned}
V_t(k,\zeta,J,M,T) &= \max_{c,l,\mu} \quad u(c,l) + \beta\mathbb{E}[V_{t+1}(k^+,\zeta^+,J^+,M^+,T^+)] \\
\text{s.t.} \quad k^+ &= (1-\delta)k + \Omega_t(1-\Lambda_t)f(k,l,\zeta,t) - c, \\
M^+ &= \Phi^M M + (E_t,0,0)^\top, \\
T^+ &= \Phi^T T + (\xi_1 F_t, 0)^\top, \\
\zeta^+ &= g^\zeta(\zeta,\omega^\zeta), \\
J^+ &= g^J(J,\omega^J)
\end{aligned}$$

Application: Carbon Tax vs. Cap-and-Trade

Policy Alternatives

- Carbon tax
- Cap-and-Trade

DSICE was used to compute optimal comovement of carbon tax and permissible emissions

- Explicitly takes into account unpredictability in economic activity
- Optimal policy is a permit supply curve with elasticity between one and two; a strong version of a price cap

Future Directions for Dynamic Programming

New tools:

- There is no curse of dimensionality in either quadrature or approximation for smooth functions - Griebel and Wozniakowski
- Massive parallelization is current supercomputer architecture; DP fits it well

Economics and OR

- The traditional ties died in late 1970's
- Economics is now hostile to introduction of OR and applied math tools; actively suppresses research that does not make economists look good
- "Soon economists will be so far behind that they will not be able to catch up"
- Hopefully concerted efforts by economists and OR researchers will prevent this