Asset Allocation using Particle Methods

John R. Birge
The University of Chicago Booth School of Business
Themes

• General SP methods are subject to the curse of dimensionality
• Randomization combined with particle methods can overcome some dimensionality effects
• With sufficient symmetry, the results for dynamic asset allocation can be effective
• Convergence results for stationary distributions
Outline

• Basic problem
• Problems of estimation
• Particle filters and sequential methods for simulation
• Results for LQG approximations and preliminary results for portfolios
• Conclusions
General Problem:

Find a set of controls, \( u_t, t = 0, \ldots, T \), to

\[
\min_{u_t \in U_t \cap A_t; x_{t+1} = g_t(x_t, u_t)} \mathbb{E} \left[ \sum_{t=0}^{T-1} f_t(x_{t+1}, u_t) \right]
\]

where \( A_t \) is \( \Sigma_t \)-measurable (nonanticipative),
\( U_t \) represents the possibility of additional constraints on the controls,
\( g_t \) represents state dynamics,
\( f_t \) represents the cost of the current action and next state transition.

Challenges:
Rapid expansion of states in dimension of \( x_t \) in each period,
Exponential growth of trees over time.
Asset Allocation Specification

Decisions and Parameters:

\( x_t(i), i = 1, \ldots, n \) - holdings in \( i \) at \( t \)
\( s_t(i), i = 1, \ldots, n \) - sales of \( i \) at \( t \)
\( b_t(i), i = 1, \ldots, n \) - purchases of \( i \) at \( t \) \((u_t = (s_t, b_t))\)
\( \tau(i), i = 1, \ldots, n \) - proportional transaction cost of trades in \( i \) at \( t \)
\( r_t(i), i = 1, \ldots, n \) - return of \( i \) at \( t \)

Objective:

\( f_T(x_{T+1}) = -u(e^T(I + \text{diag} \,(r_T))x_T) \) - (negative) utility of final wealth
\( f_t(x_{t+1}|x_t, u_t = (s_t, b_t)) = 0 \) if feasible; \(+\infty\) if not feasible for \( t < T \)

Value function: \( V_t(x_t) = \inf_{u_t} E[f_t(x_{t+1}|x_t, u_t) + V_{t+1}(x_{t+1})|x_t, u_t] \)

Constraints:

\((I + \text{diag} \,(r_t))x_t - (I + \text{diag} \,\tau)s_t + (I - \text{diag} \,\tau)b_t = x_{t+1}\)
\(e^Ts_t - e^Tb_t = 0; \ s_t, b_t \geq 0\)
Particle Methods

- Idea: Keep a limited number ($N$) of particles in each period
- Propagate forward to obtain new points
- Re-sample (re-weight) backwards to get consistent weights
- Re-balance if weights are low

Bayesian View of Particle Methods

• Suppose a distribution on $x_0$: $f(x_0)$

• Assume some transition: $p(x_t|x_{t-1})$

• Observations: $q(y_t|x_t)$

• Goal: find $P(x_t|y_1,\ldots,y_t (\ldots y_T))$

• Idea: Use Bayes Rule to find the conditional probability:

$$P(x_t|y^t) = P(x_t,y^t)/P(y^t)$$

$$\propto \prod_s q(y_s|x_s)p(x_s|x_{s-1})f(x_0)$$
Sequential Importance Resampling (SIR) Method

1. **Propagate.** Draw $x_{t+1,j}, j = 1, \ldots, N$ from $f(x_{t+1,j} | x_t,j)$.

2. **Re-weight.** Update weights, $\hat{w}_{t,j}, j = 1, \ldots, N$, to $\hat{w}_{t+1,j} = \frac{w_{t,j} g(y_{t+1} | x_{t+1})}{\sum_{j=1}^{N} \hat{w}_{t,j} g(y_{t+1} | x_{t+1})}$ for $j = 1, \ldots, N$.

3. **Re-sample.** If a re-sampling condition is satisfied, draw $N$ new samples $x'_{t+1,j}, j = 1, \ldots, N$ from $x_{t+1,j}, j = 1, \ldots, N$ with probability distribution given by $\hat{w}_{t+1,j}, j = 1, \ldots, N$; let $x_{t+1,j} = x'_{t+1,j}$ and $\hat{w}_{t+1,j} = 1/N$ for $j = 1, \ldots, N$. 
Comparison to MCMC/Particle Filter

• Idea: construct a sample distribution that fits the data using only information about one-step transitions

• Goal: maintain a constant number of samples across time instead of explosively growing tree

• Can it work? Why?
Method Structure

Propagate:
• $w_t^1, x_t^1$
• $w_t^2, x_t^2$
• ...
• $w_t^N, x_t^N$

Re-sample:
• $w_t^{'1}$
• $w_t^{'2}$
• ...
• $w_t^{'N}$
Cum. Results

- $N=50$
- Compare to Kalman Cumulative
Particle Frequency

- $N=50$
- Particle (+)
- Kalman (*)
General Optimization Extension

Problem: find $x$ to

$$\min_{x} f(x) \text{subject to } g(x) \leq 0$$

Pincus (1968): enough to sample from:

$$\pi_\kappa(x) \propto \exp\{-\kappa(f(x))\}$$

with constraints (e.g., Geman/Geman (1985)):

$$\pi_{\kappa,\lambda}(x) \propto \exp\{-\kappa(f(x) + \lambda g(x))\}$$

Extension: Augment with latent variables $(\lambda, \omega)$ then we can write

$$\pi_{\kappa,\lambda}(x) \propto \exp\{-\kappa(f(x) + \lambda g(x))\}$$

$$= e^{ax} \mathcal{E}_\omega\{e^{-\frac{x^2}{2}\omega}\} e^{bx} \mathcal{E}_\lambda\{e^{-\frac{x^2}{2}\lambda}\}$$

where the parameters $(a, b)$ depend on $(\kappa, \lambda)$ and the functional forms of $(f(x), g(x))$. 

© JRBirge  

ICSP Buzios, June 2016
Dynamic Model Extensions

- **General Idea:**

  Create a distribution over states and values (and possibly decision variables)
  
  (Use MCMC to sample – with sufficient number of particles or “annealing parameter” – marginal mean on optimum)

- **Extension for dynamics:**

  Keep constant sample particle numbers
  
  Convergence in fixed number $N$ of particles per stage to stationary distribution that is converges to optimum in $N$
General Method Structure

- $V^1_t, x^1_t, u^1_t$
- $V^2_t, x^2_t, u^2_t$
- $\ldots$
- $V^N_t, x^N_t, u^N_t$

Propagate

Optimize:

- $V^1_{t+1}, x^1_{t+1}, u^1_{t+1}, V^1_t$
- $V^2_{t+1}, x^2_{t+1}, u^2_{t+1}, V^2_t$
- $\ldots$
- $V^N_{t+1}, x^N_{t+1}, u^N_{t+1}, V^N_t$
Basic Algorithm

Initialization:
For each state $x^i_t$, $i = 1, \ldots, N$, choose $u^i_t$, $i = 1, \ldots, N$ and generate $x^i_{t+1}$, $i = 1, \ldots, N$ for $t = 1, \ldots, T$.
For each $x^i_{T+1}$, $i = 1, \ldots, N$, find value function, $V^i_{T+1}$, $i = 1, \ldots, N$.

Backward Recursion:
For each $x^i_t$, $i = 1, \ldots, N$, find $\hat{u}^i_t$, $i = 1, \ldots, N$ to solve
\[
\hat{V}^i_t = \min_{\hat{u}^i_t \in U_t(x^i_t)} \sum_{j=1}^N \frac{p(x^j_{t+1} | x^i_t, \hat{u}^i_t)}{p(x^j_{t+1} | x^j_t, u^j_t)} \frac{V^j_{t+1}}{N}.
\]
Update $u^i_t = \hat{u}^i_t$, $V^i_t = \hat{V}^i_t$; $i = 1, \ldots, N$.

Forward Recursion:
For each $x^i_t$, $i = 1, \ldots, N$ and $u^i_t$, $i = 1, \ldots, N$, generate $x^i_{t+1}$, $i = 1, \ldots, N$.

Stationary Distribution: $(x^i_t, V^i_t)$, $i = 1, \ldots, N \to (x^*_t, V^*_t)$ in distribution for each $t$ as $N \to \infty$. 
Interpretation as Importance Sampling

Basic Approach:

\[
V_t(x_t^i) = \min_{u_t^i} V_t(x_t, u_t) = \int p(x_{t+1}|x_t^i, u_t^i) V_{t+1}(x_{t+1}) dx_{t+1}
\]

\[
= \int p(x_{t+1}|x_t', u_t') \left( \frac{p(x_{t+1}|x_t^i, u_t^i)}{p(x_{t+1}|x_t', u_t')} V_{t+1}(x_{t+1}) \right) dx_{t+1}
\]

\[
\approx \sum_{j=1}^{N} \left( \frac{p(x_{t+1}^j|x_t^i, u_t^i)}{p(x_{t+1}^j|x_t^j, u_t^j)} \left( \hat{V}_{t+1}(x_{t+1}^j)/N \right) \right)
\]

where \(x_{t+1}^j \sim p(x_{t+1}|x_t^j, u_t^j)\) and \(\hat{V}_{t+1}\) is the current value function estimate at \(t+1\).

Symmetry Result: If minimization of the sample approximation is unbiased (as, e.g., in LQG problems), then the first-period solution distribution \(\hat{x}_1\) converges in mode to \(x_1^*\) for any \(N\).
Reducing Bias with MCMC

**General Idea:** Draw from the distribution of \( p(x_{t+1}|x_t, u_t^*(x_t))V_{t+1}(x_{t+1}) \).

From importance sampling:

\[
\int \frac{p(x_{t+1}|x_t, u_t^*)V_{t+1}(x_{t+1})}{\int p(x_{t+1}|x_t', u_t')V_{t+1}(x_{t+1})dx_{t+1}} \left( \int p(x_{t+1}|x_t, u_t^*)V_{t+1}(x_{t+1})dx_{t+1} \right) dx_{t+1},
\]

i.e., zero-variance samples with value: \( V_t(x_t) = \int p(x_{t+1}|x_t, u_t^*)V_{t+1}(x_{t+1})dx_{t+1} \).

**Importance Sampling version:**

\[
V_t(x_t, u_t) = \int p(x_{t+1}|x_t, u_t)V_{t+1}(x_{t+1})dx_{t+1}
\]

\[
= \int \frac{p(x_{t+1}|x_t', u_t')V_{t+1}(x_{t+1})}{\int p(x_{t+1}|x_t', u_t')V_{t+1}(x_{t+1})dx_{t+1}} \left( \frac{p(x_{t+1}|x_t', u_t')}{p(x_{t+1}|x_t', u_t')} \int p(x_{t+1}|x_t', u_t')V_{t+1}(x_{t+1})dx_{t+1} \right) dx_{t+1}
\]

\[
\approx \sum_{j=1}^N \left( \frac{p(x_{t+1}^j|x_t, u_t)}{p(x_{t+1}^j|x_t', u_t')} \right) (\hat{V}_t(x_t^j)/N)
\]

where \( x_{t+1}^j \sim p(x_{t+1}|x_t^j, u_t^j)\hat{V}_{t+1}(x_{t+1}) \) (which can be done with Metropolis-Hastings sampling).
Example: Linear-Quadratic Gaussian

• All distributions are available in closed form

\[ x_t = Ax_{t-1} + B_t u_{t-1} + v_t \]

• All normal distributions with risk-sensitive objective:

\[ -\log E(\exp(-\frac{\theta}{2}(x_2^T Q x_2 + u_1^T R u_1))) \]

where \( u_1 \) is control in the first period

Note: Equivalent to robust formulation

• Now, all can be found in closed form
Results for LQG

• Mean of $u$: $m_{u_1}$
• $m_{u_1} = L_1(\theta) x_1$, where $L_1$ is the result from the Kalman filter
• Variance of $u$: $V_u(N)$ where $V_{u_1}(N) \rightarrow 0$ as $N \rightarrow \infty$
• $m_{u_1}(\theta) \rightarrow u_1^*$ as $\theta \rightarrow 0$

where $u_1^*$ is the risk-neutral solution.
Two-Stage: Compare Direct MC

Low Risk Aversion  High Risk Aversion
Three Stage Example
Portfolio Convergence Results
Conclusions

• Techniques from Particle Filtering and MCMC can be used effectively for optimizing to reduce dimension effects
• Results work well in situations with sufficient symmetry
• Convergence in mode for problems such as asset allocation
• Many opportunities for acceleration and generalization
Thank you!

• Questions?