Quasi-Convex Stochastic Dynamic Programming

John R. Birge
University of Chicago
Booth School of Business
General Theme

- Many dynamic optimization problems dealing with commodity storage management have a value function with quasi-convex (or quasi-concave) properties.
- Taking advantage of the structure can help overcome curse of dimensionality.
- An approximation algorithm can repeatedly identify sub-level sets to obtain convergence.
Outline

• Motivation
• Basic properties
• Level-set approximation algorithm
• Computational examples
• Conclusions and next steps
Motivation

• Commodity storage characterized by inventory \((x)\) and price \((p)\): \(px\)

• For a one-stage objective \(x \geq 0\) and \(p \geq 0\), \(V(p,x) = px\) is quasi-concave:
Challenges

• Quasi-convexity is not generally additive
• Many DP objectives are additive but others are not, e.g., risk-sensitive objective:
  \[ \log \left[ \mathbb{E} \left( \exp \left( \theta \sum v_t (x_t, p_t) \right) \right) \right] \]
• When does the property persist in the value function across time?
• How to take advantage of the property when present?
Simple Example: Commodity Management

Objective (with action $u$, capacity $C$, discount factor $\delta$):

$$V_t(x_t, p_t) = \max_{x_t \leq u \leq C - x_t} -p_t u + \delta E[V_{t+1}(x_t + u, p_{t+1})]$$

- Quasi-concave at end of horizon (T)
- Induction possible?
- Assume: $(p_{t+1}/p_t) \sim F$ (fixed stationary)
Quasi-Concavity Result for $V_t$

- Basic induction:

$V_t$ is then quasi-concave if

$$V_t(x, p) \geq \min\{V_t(x_1, p_1), V_t(x_2, p_2)\},$$

for any such combination of points, $(x_1, p_1)$ and $(x_2, p_2)$. Without loss of generality, suppose $x_1 < x_2$ and let

$$x = x_1 + \Delta = x_2 - \Delta,$$

and

$$p = U p_1 = D p_2.$$  \hspace{1cm} (2)

Suppose

$$V_t(x_1, p_1) = -u_1 p_1 + E_{p_1} V_{t+1}(x_1 + u_1, p'),$$

and

$$V_t(x_2, p_2) = -u_2 p_2 + E_{p_2} V_{t+1}(x_2 + u_2, p').$$  \hspace{1cm} (3)

\hspace{1cm} (4)

\hspace{1cm} (5)
Quasi-concavity of Value Function

Suppose policy $\sigma_1$ starting from $(x, p)$ at stage that sets $u_1^\Delta = u_1 + \Delta$ at stage $t$ and then follows policy $\pi_1(\omega)$ (optimal along all $p_t(\omega) = w_t(\omega)p_1$ at time $t$; $\sigma_2$ starting from $(x, p)$ with $u_2^\Delta = u_2 - \Delta$ at $t$ and then $\pi_2(\omega)$-optimal for $p_t(\omega) = v_t(\omega)p_2$; $\sigma$ be the policy that randomly chooses $\sigma_1$ or $\sigma_2$ from state $(x, p) \to$

$$V_t(x, p) \geq V_{t, \sigma}(x, p) = \frac{1}{2}(- (u_1 + \Delta) U p_1 + U E p_1 V_{t+1}(x_1 + u_1, p')$$

$$- (u_2 - \Delta) D p_2 + D E p_2 V_{t+1}(x_2 + u_2, p'))$$

$$= \Delta(- U p_1 + D p_2) + \frac{1}{2}(U V_t(x_1, p_1) + D V_t(x_2, p_2))$$

$$\geq \frac{1}{2}(U + D) \min(V_t(x_1, p_1), V_t(x_2, p_2))$$

$$= \frac{p_1^2 + p_2^2 + 2p_1 p_2}{4p_1 p_2} \min(V_t(x_1, p_1), V_t(x_2, p_2)),$$

$$\geq \min\{V_t(x_1, p_1), V_t(x_2, p_2)\}$$

using $p = U p_1 = D p_2$, $U = \frac{p_1 + p_2}{2p_1}$, $D = \frac{p_1 + p_2}{2p_2}$, and that $2p_1 p_2 \leq p_1^2 + p_2^2$. 

© JRBirge  
SIAM FM12, MSP, 10 July 2012
Analysis and Extensions of Result

• Extends to proportional transaction costs directly

• Key assumption: Price process increments are independent of the price level (as in a log-normal/GBM model)

• Extends to other price processes as well:

  Distributions such that $V_{t,\pi_1}(x_1, U p_1) = U' V_t(x_1, p_1)$ and $V_{t,\pi_1 2}(x_2, D p_2) = D' V_t(x_1, s_1)$ where $U' + D' \geq 2$. 
Extended Processes

- **Ornstein-Uhlenbeck:**

\[
d \log p_t = \theta(\mu - \log p_t)dt + \sigma dW_t, \tag{1}
\]

where \(\theta, \mu,\) and \(\sigma\) are scalars; \(W_t\) is a standard Brownian motion. The price \(p_t\) given initial value \(p\) is then

\[
p_t = \exp(\log pe^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(q-t)} dW_q). \tag{2}
\]

Let \(p = Up_1 = Dp_2, p^i_t\) be spot price at \(t\) given spot price \(p_i\) at 0, \(i = 1, 2,\) then

\[
p^1_t \sim p_t U e^{-\theta t} = p_t U', p^2_t \sim p_t D e^{-\theta t} = p_t D', \tag{3}
\]

the condition, \(U' + D' \geq 2\) holds for sufficiently small \(\theta\) and sufficiently large time intervals.
Algorithm Motivation

• Assume: Quasi-concave (-convex) structure
• Basic idea: approximate the level sets of the value function at each stage
• Use the level-set approximation in stage t+1 to construct an approximation at t.
• Use outer linearization to construct each successive approximation
Quasi-concave Dynamic Program Method (QDPM)

1. Set $U_t(x, p)$ for all $t$. Let $U_T(x, p) = V_T(x, p)$. Set $t = T - 1$.

2. REPEAT:
   
   (a) Randomly select $(x_t, p_t)$;
   (b) Find $\hat{U}_t(x_t, p_t) = \max_{u \in A(x_t, p_t)} v_t(u, x_t, p_t) + \sum_{z \in Z_{t+1}(u, x_t, s_t)} \pi_{t+1}(z, a, x_t, p_t) U_{t+1}(z)$;
   (c) Perform Bound Update;
   (d) Choose $t$.

3. UNTIL $U_1(x_1, p_1)$ is stable.

Bound Update

Construct upper bound on the value function using outer linearization:

$$U_t(x, p) \leftarrow \min(U_t(x, p), \hat{U}_t(x, p))$$  \hspace{1cm} (1)

for all $(x, p)$ such that $\nabla \hat{U}_t(x_t, p_t)^T (x - x_t, p - p_t) \leq 0$. 
Bound Update

\[ \nabla \hat{U}_t(x_t, p_t)(x - x_t, p - p_t) \leq 0 \]

\[ U_t(x, p) \leq \hat{U}(x_t, p_t) \]
Examples

Setup: Lognormal prices (log-mean: 1, log-st.dev.: 0.4)
Transaction cost: 5% proportional

Note: bound update must achieve sufficient accuracy to also maintain quasi-concavity
Use of 1000, 500, and 250 random samples per period
Approximations at $T-1$

Figure 1: Contour plots of value function and upper bounds for prices in [0,3.5] and inventory in [0,1] one stage before the end of horizon.
Upper Bounds at $T-5$

Figure 2: Contour plots of value function and upper bound for prices in $[1,3.5]$ and inventory in $[0,1]$ five stages before the end of horizon.
Observations from Example

For the $T-1$-stage results:
Close to original for $N_{T-1}=1000$
More differences for $N_{T-1}=500$ and $N_{T-1}=250$
increasingly less consistent in shape with fewer samples.

For the $T-5$-stage results: Similar pattern – more error – less consistent level sets with
$N_{T-5}=250$
Conclusions

• Quasi-convexity may persist in some commodity inventory control problems

• Possible resolutions with level set approximations

• Limitations:
  • Accuracy needed in the approximation to preserve quasi-concavity
  • Unclear results for higher-dimensions
Future Work

• Find other structures (such as risk-sensitive objective) with quasi-convex structure
• Identify methods to maintain the property in the approximation
• Extend results for multiple dimensions
Thank you!