**Introduction**

Motivation: Multi-stage decision making under uncertainty:
- Utilize historical observations directly in the solution procedure
- Handle data sparsity
- Two-stage setting: Exploit side information in decision making
- Dynamic setting: Mitigate the three curses of dimensionality

Our solution:
- Use Nadaraya-Watson (NW) kernel regression to estimate conditional expectations
- Use robust optimization to reduce the effects of estimation errors

**Stochastic Programming (SP)**

Stochastic program with side information: SP: \( \min_{x \in X} \{ c(x, \xi); \xi - \xi_0 \} \)

Example: News vendor problem
- The current demand \( \xi_t \) is uncertain but depends on the demand \( \xi_0 \) of the period before.
- The conditional distribution of \( \xi_1 \) given \( \xi_0 \) is less dispersed than the marginal distribution of \( \xi_1 \).

| \( \xi_1 \): | \( \xi_1 = 0 \): | \( \xi_1 = 10 \), var\( (\xi_1) = 1 \) |
| \( \xi_0 \): | \( \xi_0 - \xi_0 = 0 \): | \( \xi_0 = 0 \), var\( (\xi_0 - \xi_0) = 0 \) |

Statistics: \( \{ \xi_1, \xi_0 \} \) - 10th & 90th percentiles

![Distribution of \( \xi_1 \) and \( \xi_0 \)](image)

**Data-Driven SP**

Motivation:
- The joint distribution of \( \xi_1 \) and \( \xi_0 \) is unknown
- Only few historical observations \( (\xi_0, \xi_1) \) are available

Data-driven stochastic program:
- Estimate the conditional expectation using NW kernel regression:
  \[
  \text{DSP: } \min_{x \in X} \{ c(x, \xi); \xi - \xi_0 \} = \min_{x \in X} \sum_{i=1}^{N} a(x, \xi_i) \]
  \[
  a(x, \xi) = \frac{1}{N} \sum_{i=1}^{N} \rho(x, \xi_i) \]
  \[
  \rho(x, \xi) = \frac{\xi(x) - \xi(x, \xi)}{\sum_{j=1}^{N} \xi(x, \xi_j)}
  \]
- \( a(x, \xi) \) are normalized weights constructed using kernel \( K(\cdot) \)

Shortcomings:
- When few observations are available the NW estimate of the conditional expectation exhibits a high variability
- Estimation errors lead to an optimistic downward bias in DSP

| \( \xi_1 \): | \( \xi_1 = 0 \): | \( \xi_1 = 10 \), var\( (\xi_1) = 1 \) |
| \( \xi_0 \): | \( \xi_0 - \xi_0 = 0 \): | \( \xi_0 = 0 \), var\( (\xi_0 - \xi_0) = 0 \) |

Statistics: \( \{ \xi_1, \xi_0 \} \) - 10th & 90th percentiles

![Distribution of \( \xi_1 \) and \( \xi_0 \)](image)

**Dynamic Programming (DP)**

Stochastic dynamic program:
- \( V(s_t, \xi_t) = \min_{a_t} c(s_t, a_t, \xi_t) + E \left[ V(s_{t+1}, \xi_{t+1}) | s_t, a_t, \xi_t \right] \)
  \( s_t, a_t, \xi_t \in X \), \( s_{t+1} = g(s_t, a_t, \xi_t) \)
- Endogenous state: \( s_t \) is decision-dependent
- Exogenous state: \( \xi_t \) is decision-independent

Data-driven dynamic program:
- \( V(s_t, \xi_t) = \min_{a_t} c(s_t, a_t, \xi_t) + E \left[ \sum_{i=1}^{N} a_t \xi_i V(s_{t+1}, \xi_{t+1}) | s_t, a_t, \xi_t \right] \)
  \( s_t, a_t, \xi_t \in X \), \( s_{t+1} = g(s_t, a_t, \xi_t) \)

Robust data-driven dynamic program:
- \( V(s_t, \xi_t) = \min_{a_t} c(s_t, a_t, \xi_t) + \max_{p \in P} E \left[ \sum_{i=1}^{N} \rho_i V(s_{t+1}, \xi_{t+1}) | s_t, a_t, \xi_t \right] \)
  \( s_t, a_t, \xi_t \in X \), \( s_{t+1} = g(s_t, a_t, \xi_t) \)

Separate architectures for endogenous and exogenous states:
- Endogenous: conic representable approximatively
- Exogenous: NW kernel regression to estimate the conditional expectation

Complexity of RDDP:
- If \( c \) is convex quadratic \( \Rightarrow g \) is affine \( \Rightarrow x_t \) is second-order conic representable, then RDDP reduces to a second-order cone program (SOCP) that can be solved in \( O(\sqrt{N}m) \) iterations, where \( m \) = \#decision variables

**Algorithm**

Inputs: State trajectories \( s_t^k \) \( 1 \leq k \leq K \), observation histories \( \xi_t^k \) \( 1 \leq i \leq N \)
- For \( t=1,...,T \)
  - For \( i=1,..,N \)
    - Evaluate RDDP with inputs \( V_t(\xi_t^k) \) in the endogenous state
    - Construct parametric approximation \( V_t(\xi_t^k) \) for \( i = 1,...,N \) and \( t = 1,...,T \)

Main features of algorithm:
- Number of RDDP problems to solve is \( O(NKT) \)
- The NK optimization problems are independent (parallelizable)

**Results**

Index tracking:
- Track S&P 500 Index over 20 trading days using a portfolio of indices: NASDAQ, Russell 2000, S&P MidCap 400, and AMEX Major Market
- Test data: 16.3.1998 - 3.8.2013 (192 months)
- RDDP outperforms DDP and LSPI substantially
- The distribution of tracking errors generated by RDDP stochastically dominates those generated by DDP and LSPI

<table>
<thead>
<tr>
<th>Statistic</th>
<th>LSPI</th>
<th>DDP</th>
<th>RDDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.99</td>
<td>4.70</td>
<td>1.28</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>11.70</td>
<td>15.07</td>
<td>2.23</td>
</tr>
<tr>
<td>90th prct.</td>
<td>14.60</td>
<td>9.05</td>
<td>2.85</td>
</tr>
<tr>
<td>Worst case</td>
<td>126.71</td>
<td>157.20</td>
<td>18.83</td>
</tr>
</tbody>
</table>

![CDF of tracking errors](image)