Information, Noise and Lies: The Evolving Discovery of Misinformation in Rail Transportation

Tenth International Conference on Stochastic Programming
University of Arizona
October, 2004

Warren Powell
CASTLE Laboratory
Princeton University
http://www.castlelab.princeton.edu

© 2003 Warren B. Powell, Princeton University
Option 1: Send directly to customers
Option 1: Send directly to customers
Option 2: Send to regional depots
Option 1: Send directly to customers
Option 2: Send to regional depots
Option 3: Send to classification yards
The importance of uncertainty

Available resource:

Possible customer demand

Possible customer demand
The importance of uncertainty

- Why do stochastic models work?
  - The newsvendor effect:

  - Revenue from an additional boxcar.
  - Cost of an additional car.

  Point forecast

Demand

0

5

$100
The importance of uncertainty

Why do stochastic models work?

» The newsvendor effect:

Revenue from an additional boxcar.

Cost of an additional car.
The car cycle

- Empty movement
- Load car
- Loaded movement
- Unload car
The planning process
The planning process
The planning process
A myopic model
If a customer releases a car in the beginning of the week, you have to send it somewhere right away. We can’t hold cars, wait for an order, and then send them.
The planning process
Forecasting

Actual vs. predicted car demands

© 2004 Warren B. Powell
Forecasting

Actual vs. predicted car supplies
Two-stage problems

- Two-stage resource allocation under uncertainty
Two-stage problems

- If we model different regions as separable....
Optimization frameworks

- We obtain piecewise linear recourse functions for each region.
Optimization frameworks

- The function is piecewise linear on the integers.
Optimization frameworks

- Using standard network transformation:

Each link captures the marginal reward of an additional car.
Two-stage problems
Literature review

■ Piecewise linear bounds:

1987: Wallace, S.W., “A Piecewise Linear Upper Bound on the Network Recourse Function”

■ Static piecewise linear approximations:

Literature review

Static piecewise linear approximations (cont’d)


Literature review

- Adaptive linear approximations

1998: Powell, W.B. and T. Carvalho, “Dynamic Control of Logistics Queueing Network for Large-Scale Fleet Management”
Literature review

- Adaptive nonlinear/piecewise linear approximations


2001: Godfrey, G. and W. B. Powell, “An Adaptive, Distribution-Free Approximation for the Newsvendor Problem with Censored Demands, with Applications to Inventory and Distribution Problems”


Two-stage problems
Two-stage problems

We estimate the functions by sampling from our distributions.

Marginal value:

\[ \hat{v}_1(\omega^n) \rightarrow R_1^n \rightarrow D_1(\omega^n) \]
\[ \hat{v}_2(\omega^n) \rightarrow R_2^n \rightarrow D_2(\omega^n) \]
\[ \hat{v}_3(\omega^n) \rightarrow R_3^n \rightarrow D_3(\omega^n) \]
\[ \vdots \]
\[ \hat{v}_5(\omega^n) \rightarrow R_5^n \rightarrow D_C(\omega^n) \]
Two-stage problems

\[ \begin{align*}
\begin{array}{c}
\vphantom{v^+} \\
\vphantom{v^-}
\end{array}
\end{align*} \right\} \text{Duals}

\begin{align*}
\text{Profits} & \quad v^- \\
\text{Number of vehicles} & \quad v^+
\end{align*}
Two-stage problems

- Left and right derivatives are used to build up a nonlinear approximation of the subproblem.
Two-stage problems

- Left and right derivatives are used to build up a nonlinear approximation of the subproblem.
Two-stage problems

- Each iteration adds new segments, as well as refining old ones.
Two-stage problems

![Graph showing the relationship between number of resources and approximate value function. The graph displays different iterations (1 Iter, 2 Iter, 5 Iter, 10 Iter, 15 Iter, 20 Iter) with corresponding approximate value functions. The x-axis represents the number of resources, and the y-axis represents the approximate value function. The graph includes the exact value function for comparison.]
Two-stage problems

- Two-stage stochastic programming.
Two-stage problems

- Exact solutions using Benders:
  - “L-Shaped” decomposition (Van Slyke and Wets)
  - Stochastic decomposition (Higle and Sen)
  - CUPPS (Chen and Powell)
Two-stage problems

Variations on Bender’s decomposition

Iterations
Two-stage problems

Variations on Bender’s decomposition

Separable approximation

Iterations
Two-stage problems

Variations on Bender’s decomposition

Separable approximation

Deterministic approximation

Iterations

© 2004 Warren B. Powell
We have different box car types, and some customers are flexible about which types we send them.
Two-stage problems

- Two-stage resource allocation under uncertainty
Two-stage problems

- We normally think of sending cars to locations…

© 2004 Warren B. Powell
Two-stage problems

The set of car types.
Two-stage problems

The set of locations
Two-stage problems

Geographical substitution
Two-stage problems

Functional substitution
Two-stage problems

... actually, we “create” cars with different attributes.
Multiattribute resources

- Assets can have a number of attributes:

\[
a = \begin{bmatrix}
\text{Location} \\
\text{Location} \\
\text{Boxcar type}
\end{bmatrix}
\]

\[|\mathcal{A}| \approx 200 \quad 2000\]
A car arriving on Thursday can serve an order that was available on Tuesday. We should penalize this, but we can do it.
Empty movement

Loaded movement
Empty movement

Loaded movement
Multiperiod travel times

Available:

\[
t \quad t+1 \quad t+2 \quad t+3
\]
Multiperiod travel times

Available:

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>t+1</th>
<th>t+2</th>
<th>t+3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2004 Warren B. Powell
Multiperiod travel times

The set of locations
Multiperiod travel times

Available now
Available t+1
Available t+2
Available t+2
Multiattribute resources

- If the longest trip is seven days...

\[
\begin{bmatrix}
\text{Location} \\
\text{Boxcar type}
\end{bmatrix} =
\begin{bmatrix}
\text{Location} \\
\text{Boxcar type} \\
\text{Time to dest.}
\end{bmatrix}
\]

\[|A| \approx 200 \quad 2000 \quad 14,000\]
But the travel times are not that long! We can load and unload cars multiple times within two weeks.
Multistage problems
Multistage problems
Multistage problems
Multiattribute resources

For multiperiod problems, we have to capture the time period:

\[ a = \begin{bmatrix}
\text{Location} & \text{Location} & \text{Location} \\
\text{Boxcar type} & \text{Boxcar type} & \text{Time to dest.}
\end{bmatrix} \]

\[ |A| \approx 200 \quad 2000 \quad 14,000 \]
Multiattribute resources

For multiperiod problems, we have to capture the time period:

\[ a = \begin{bmatrix} \text{Time} \\ \text{Location} \\ \text{Boxcar type} \end{bmatrix} \begin{bmatrix} \text{Time} \\ \text{Location} \end{bmatrix} \begin{bmatrix} \text{Time} \\ \text{Location} \end{bmatrix} \begin{bmatrix} \text{Time to dest.} \end{bmatrix} \]

\[ |A| \approx 4,000 \quad 40,000 \quad 280,000 \]
That’s great! Let’s see how it works on our past history!

History!!?? Watch out! History is deterministic!
A single commodity flow problem

- A pure (deterministic) network:
Number of iterations

Percent of LP optimal

0 20 40 60 80 100 120 140 160 180 200

70% 75% 80% 85% 90% 95% 100%

Number of iterations of CAVE training on expectation percentage of network optimal solution.
Multicommodity flow
The optimal solution (continuous) Using separable piecewise linear approximations (integer)
Wait a minute! When the shipper books an order, he doesn’t tell us where it is going!
Multiatribute resources

Information at time $t$

Actionable time

Locations
Multiattribute resources

Information at time $t+1$

Actionable time

Locations

1

2

3

4

5
Multiattribute resources

Information at time $t+2$

Actionable time

Locations

© 2004 Warren B. Powell
Multiattribute resources

Information at time t+2

Actionable time

Locations
You are assuming that the travel times between points takes an integer number of days. This is just not realistic!
Multistage problems

- We had assumed that the time discretization of the information process matched the physical process.
Multistage problems

- If we want finer time steps, we had to run more iterations:
Multiattribute resources

- With one day time steps:

\[
a = \begin{bmatrix}
    \text{Time} & \text{Time} & \text{Time} \\
    \text{Location} & \text{Location} & \text{Location} \\
    \text{Boxcar type} & \text{Boxcar type} & \text{Time to dest.}
\end{bmatrix}
\]

\[
|\mathcal{A}| \approx 4,000 \quad 40,000 \quad 280,000
\]
Multiattribute resources

- With four hour time steps for informational and physical processes:

\[
a = \begin{bmatrix} 
\text{Time} & \text{Time} & \text{Time} \\
\text{Location} & \text{Location} & \text{Location} \\
\text{Boxcar type} & \text{Boxcar type} & \text{Time to dest.}
\end{bmatrix}
\]

\[|\mathcal{A}| \approx 24,000 \quad 240,000 \quad 40,080,000\]
Multiattribute resources

- With four hour time steps for just physical processes:

\[ a = \begin{bmatrix} 
\text{Time} \\
\text{Location} \\
\text{Boxcar type} \\
\text{Time to dest.} 
\end{bmatrix} \]

\[ |A| \approx 4,000 \quad 40,000 \quad 1,680,000 \]
But the cars break down… we have to model the repair status so we capture the right number of usable cars.
Repair status

Acceptable  Minor repair  Major repair

© 2004 Warren B. Powell
In practice, there are a number of parallel information processes taking place:
Multiattribute resources

- When we add in repair status:

\[
a = \begin{bmatrix}
\text{Time} \\
\text{Location} \\
\text{Boxcar type} \\
\text{Time to dest.}
\end{bmatrix}
\]

\[
|A| \approx 4,000 \quad 40,000 \quad 1,680,000
\]
Multiattribute resources

When we add in repair status:

\[ a = \begin{bmatrix}
\text{Time} & \text{Time} & \text{Time} & \text{Time} \\
\text{Location} & \text{Location} & \text{Location} & \text{Location} \\
\text{Boxcar type} & \text{Boxcar type} & \text{Boxcar type} & \text{Boxcar type} \\
\text{Time to dest.} & \text{Time to dest.} & \text{Time to dest.} & \text{Repair status}
\end{bmatrix} \]

\[ |A| \approx 4,000 \quad 40,000 \quad 1,680,000 \quad 5,040,000 \]
Two-stage problems

- We need to compute duals for all the attributes...

Marginal value:

\[ \hat{v}_1(\omega^n) \quad R_1^n \rightarrow \]

\[ \hat{v}_2(\omega^n) \quad R_2^n \rightarrow \]

\[ \hat{v}_3(\omega^n) \quad R_3^n \rightarrow \]

\[ \hat{v}_4(\omega^n) \quad R_4^n \rightarrow \]

\[ \hat{v}_5(\omega^n) \quad R_5^n \rightarrow \]

\[ D_1(\omega^n) \]

\[ D_2(\omega^n) \]

\[ D_3(\omega^n) \]

\[ \ldots \]

\[ D_C(\omega^n) \]
Two-stage problems

… even the ones with zero flow.

Marginal value:

\[ \hat{v}_1(\omega^n) \ R_1^n > 0 \rightarrow \]

\[ \hat{v}_2(\omega^n) \ R_2^n > 0 \rightarrow \]

\[ \hat{v}_3(\omega^n) \ R_3^n = 0 \rightarrow \]

\[ \hat{v}_4(\omega^n) \ R_4^n = 0 \rightarrow \]

\[ \hat{v}_5(\omega^n) \ R_5^n = 0 \rightarrow \]
Optimization frameworks

Attribute space is now too large to enumerate.
The biggest computational expense is simply generating the links out of the nodes.
Aggregation

\[ \bar{v}_{NE} \]

NE region

PA

\[ \nu_{PA} = ? \]

TX

© 2004 Warren B. Powell
\hat{v}(\omega)
$\overline{V}_a^{-1} = \frac{\hat{v}_1 + \hat{v}_2 + \hat{v}_3}{3}$
Aggregation

Number of Iterations

Percent of Optimal

1 x 1 grid
Aggregation

![Graph showing percent of optimal solution over iterations for different grid sizes. The graph includes lines for 5x5 Grid, 10x10 Grid, 20x20 Grid, 1x1 Grid, and 2x2 Grid. The 2x2 Grid line is highlighted in magenta.](image)

© 2004 Warren B. Powell
Aggregation

Number of Iterations

Percent of Optimal

5 x 5 grid

10x10 Grid
20x20 Grid
1x1 Grid
2x2 Grid
5x5 Grid
Aggregation

Number of Iterations

Percent of Optimal

10 x 10 grid

20x20 Grid
1x1 Grid
2x2 Grid
5x5 Grid
10x10 Grid
Multiattribute resources

- We can define a family of aggregation functions:

\[
G^3(a) \quad G^2(a) \quad G^1(a) \quad G^0(a)
\]

\[
a = \begin{bmatrix}
\text{Time} & \text{Time} & \text{Time} & \text{Time} \\
\text{Location} & \text{Location} & \text{Location} & \text{Location} \\
\text{Boxcar type} & \text{Boxcar type} & \text{Boxcar type} & \text{Boxcar type} \\
\text{Time to dest.} & \text{Time to dest.} & \text{Time to dest.} & \text{Time to dest.} \\
\text{Repair status} & & & \\
\end{bmatrix}
\]

\[
|A| \approx 5,040,000 \quad 1,680,000 \quad 40,000 \quad 4,000
\]
Hierarchical aggregation

- We can then use all levels of aggregation at the same time:

\[
\tilde{\nu}_a = \sum_g w_a^{(g)} \nu^{(g)}_a
\]
Hierarchical aggregation

Optimal weights change as the algorithm progresses:

- Weight on most disaggregate level
- Weight on most aggregate levels

![Graph showing weight changes over iterations and aggregation levels.](image-url)
Hierarchical aggregation

© 2004 Warren B. Powell
Hierarchical aggregation

Weighted Combination

Aggregate

Disaggregate
Flows from history

Historical flows of CN locomotives (blue)
Flows from history

Flows from the model

CN locomotives: history (blue) and LRM without patterns (red)
But we just don’t move our cars that way! If you try to send cars from Buffalo through the Cleveland yard, it comes in on the wrong tracks and can’t make the connection.
Expert knowledge

- Usual modeling approach: bottom up engineering

\[ x^* = \arg \min cx \]

Subject to: \( Ax = b, \quad x \geq 0 \)
Expert knowledge

- Bottom up/top down modeling:

**Patterns**

Specify the behaviors you want at a general level.

Specify costs, driver availability, work rules, routing preferences, load avail.

**Engineering**

© 2004 Warren B. Powell
Expert knowledge

- Pattern matching

\[ x^* = \arg \min cx + \rho H(x, \rho) \]

The “happiness” function – measures the degree to which model behavior agrees with a knowledgeable expert.

\[ H(x, \rho) = \| G(x) - \rho \| \] where \( G(x) \) is an aggregation function
Historical patterns
Expert knowledge

- Our expert:
Flows from history

Flows from the model

CN locomotives: history (blue) and LRM without patterns (red)
Flows from history

Flows from the model

CN locomotives: history (blue) and LRM with patterns (red)
You can’t trust what the customer asks for. They don’t believe we are going to give them what they need, so they ask for more than they need.
Cars ordered and the cars actually loaded

Cars ordered in advance

Cars loaded

Week Ending

- Booked Orders
- Cars Supplied
- Cars Loaded
Empty miles as a percent of total miles

History

“Optimized” without adaptive learning

“Optimized” with adaptive learning