Ellipsoidal Methods for Adaptive Choice-based Conjoint Analysis

Juan Pablo Vielma
Massachusetts Institute of Technology

March, 2016.

Joint work Denis Sauré
(Custom) Product Recommendations via CBCA

<table>
<thead>
<tr>
<th>Feature</th>
<th>SX530</th>
<th>RX100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom</td>
<td>50x</td>
<td>3.6x</td>
</tr>
<tr>
<td>Prize</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>15.68 ounces</td>
<td>7.5 ounces</td>
</tr>
<tr>
<td>Prefer</td>
<td>✔️</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>TG-4</th>
<th>G9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterproof</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Prize</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>7.36 lb</td>
<td>7.5 lb</td>
</tr>
<tr>
<td>Prefer</td>
<td>□</td>
<td>✔️</td>
</tr>
</tbody>
</table>

We recommend:

- **TX530**: $249.99, 50x Zoom, 15.68 ounces, Yes
- **RX100**: $399.99, 3.6x Zoom, 7.5 ounces

- **TG-4**: $249.99, Yes, Electronic Viewfinder
- **G9**: $399.99, No, Optical Viewfinder

Ellipsoidal Methods for Adaptive CBCA
(Custom) Product Recommendations via CBCA

• Individual preference estimates with few questions (5):
  – Need very accurate question = adaptive
  – Still need confidence measure on estimates
• Minimize uncertainty / variance good, but secondary:
  – Objective is good recommendation (M-Efficiency)
    • Final use of preference is risk-averse optimization problem
  – Need intuitive geometric model to combine learning with optimization
Towards Optimal Product Recommendation

- Find enough information about preferences to recommend

<table>
<thead>
<tr>
<th>Feature</th>
<th>SX330</th>
<th>RX100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom</td>
<td>50x</td>
<td>3.6x</td>
</tr>
<tr>
<td>Price</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>15.68 ounces</td>
<td>7.5 ounces</td>
</tr>
<tr>
<td>Prefer</td>
<td>✓</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>TG-4</th>
<th>G9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterproof</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Price</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>7.36 lb</td>
<td>7.5 lb</td>
</tr>
<tr>
<td>Prefer</td>
<td>□</td>
<td>✓</td>
</tr>
</tbody>
</table>

We recommend:

- How do I pick the next question to obtain the largest reduction of uncertainty or “variance” on preferences
# Choice-based Conjoint Analysis

<table>
<thead>
<tr>
<th>Feature</th>
<th>Chewbacca</th>
<th>BB-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wookiee</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Droid</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Blaster</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Product Profile

\[
\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x^2
\]
MNL Preference Model

- Utilities for 2 products, d features
  \[ U_1 = \beta \cdot x^1 + \epsilon_1 = \sum_{i=1}^{d} \beta_i x^1_i + \epsilon_1 \]
  \[ U_2 = \beta \cdot x^2 + \epsilon_2 = \sum_{i=1}^{d} \beta_i x^2_i + \epsilon_2 \]

- Utility maximizing customer: \( x^1 \succeq x^2 \iff U_1 \text{"\geq"} U_2 \)

- Noise can result in response error:
  \[ \mathbb{P} \left( x^1 \succeq x^2 \mid \beta \right) = \frac{e^{\beta \cdot x^1}}{e^{\beta \cdot x^1} + e^{\beta \cdot x^2}} \]
Next Question To Reduce “Variance”: Bayesian

Prior Distribution of $\beta$

- Update uses MNL response error
- Question Selection: Enumeration
- Recommendation: Risk-averse Stochastic Optimization
Next Question To Reduce “Variance”: Polyhedral

Toubia, Hauser and Simester, ‘04

Polyhedron Containing $\beta$

- Update ignores response error $\times$
- Question Selection: (Multi-Obj.) Discrete Optimization $\checkmark$
- Recommendation: Robust Optimization $\checkmark$

Ellipsoidal Methods for Adaptive CBCA
“Improving” the 2004 Polyhedral Method

• More “re-interpreting” ideas from Toubia, Hauser and Simiterer, ’04 (and Toubia, Hauser and Garcia ’07)

• Our “improvements”:

1. **Incorporate response error**
   • Adaptations by Toubia, Hauser and Garcia ’07 and Bertsimas and O’Hair ’13
     – Not MNL model
     – Loose simple geometric interpretation = complicates update, question selection and recommendation problem
   • Replace polyhedra with ellipsoids = have your cake and eat it too!

2. **“Improve” question selection**
   • Optimize widely used variance metric = D-efficiency
   • Just the right balance from guidelines from Toubia et al. 2004
Polyhedral Method
Preference Model and Geometric Interpretation

- Utilities for 2 products, d features, logit model
  \[ U_1 = \beta \cdot x^1 + c_1 = \sum_{i=1}^{d} \beta_i x_i^1 + c_1 \]
  \[ U_2 = \beta \cdot x^2 + c_2 = \sum_{i=1}^{d} \beta_i x_i^2 + c_2 \]

  part-worths
  product profile
  noise (gumbel)

- Utility maximizing customer
  - Geometric interpretation of preference for product 1 without error
    \[ x^1 \succeq x^2 \iff U_1 \succeq U_2 \]
Geometric prior for $\beta$

$\mathbf{x}^1 \succeq \mathbf{x}^2$

2nd geometric posterior for $\beta$

$\mathbf{x}^3 \succeq \mathbf{x}^4$

$\mathbf{x}^5 \succeq \mathbf{x}^6$

$\beta \cdot (x^1 - x^2) \geq 0$

$\beta \cdot (x^3 - x^4) \geq 0$

$\beta \cdot (x^5 - x^6) \geq 0$
Polyhedral: Estimation and Question Selection

Good estimator for $\beta$?

Center of ellipsoid $\beta_0$.

1. Choice balance
2. Postchoice symmetry

Ellipsoidal Methods for Adaptive CBCA
Polyhedral Method: Non-ellipsoidal Sets

Idea from Nonlinear Programming (NLP): Approximate ellipsoid through analytic center.

\[
\begin{align*}
\beta \cdot (x^1 - x^2) & \geq 0 \\
\beta \cdot (x^3 - x^4) & \geq 0 \\
\min \sum_i \log \left( \beta \cdot (x^{2i-1} - x^{2i}) \right) & \\
\end{align*}
\]
Incorporating Response Error
First Improvement: Ellipsoidal Updates

• Polyhedral updates
  – Assumes no errors
  – Region complexity increases

• NLP again: ellipsoid method
  – Use minimum volume ellipsoid = simple formula ...
  – or use corrected ellipsoid = simple modification to formula
Distributions and Credibility Ellipsoids

Prior distribution of $\beta$

$\beta \sim N(\mu, \Sigma)$

90% confidence/credibility ellipsoid

$$(\beta - \mu)' \cdot \Sigma^{-1} \cdot (\beta - \mu) \leq r$$
Answers with Error: Logit Probabilities

Likelihood Function

\[ P(x^1 \geq x^2 \mid \beta) = \frac{e^{\beta \cdot x^1}}{e^{\beta \cdot x^1} + e^{\beta \cdot x^2}} \]

Question/Answer

\[ x^1 \geq x^2 \]
Bayesian Update and Geometric Updates

Prior distribution

Prior ellipsoid

Answer likelihood

Question/Answer

Posterior distribution

Posterior ellipsoid

Ellipsoidal Methods for Adaptive CBCA
Geometric Comparison of Updates

Min. Volume Ellipsoid

Corrected Ellipsoid

Bayesian for Normal Approx.

Simple Formula

Simple Formula

1-dim integral 😊

$10^4 \rightarrow 10^7$ samples

Ellipsoidal Methods for Adaptive CBCA
Computational Comparison of Updates

- Gaussian prior and 90% credibility ellipsoid, 100 inst.
  - 12 features, 2 profiles and 5 questions

<table>
<thead>
<tr>
<th></th>
<th>Polyhedral</th>
<th>Ellipsoidal</th>
<th>Corrected Ellipsoidal</th>
<th>1-step Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasible $\beta$</td>
<td>0.53</td>
<td>1</td>
<td>1</td>
<td>0.93</td>
</tr>
<tr>
<td>Distance (scaled)</td>
<td>0.92</td>
<td>0.86</td>
<td>0.88</td>
<td>0.85</td>
</tr>
<tr>
<td>Gaussian Volume</td>
<td>0.03</td>
<td>0.85</td>
<td>0.82</td>
<td>0.40</td>
</tr>
</tbody>
</table>

“Better” measure of variance = D-Efficiency
Improving Question Selection: Optimizing D-Efficiency
D-Efficiency and Posterior Covariance Matrix

- D-Efficiency:
  \[ f(x_1, x^2) := \mathbb{E}_{\beta, x_1 \preceq x^2} \left( \det(\Sigma_i)^{1/p} \right) \]
- \( p = 2 \) proportional to expected volume of posterior ellipsoid

- Even evaluating expected D-Efficiency for a question requires multidimensional integration
Back to Question Selection: Property Trade-off

\[(\beta - \mu)' \cdot \Sigma^{-1} \cdot (\beta - \mu) \leq r\]

• Choice balance:
  – Minimize distance to center
    \[\mu \cdot (x^1 - x^2)\]

• Postchoice symmetry:
  – Maximize variance of question
    \[(x^1 - x^2)' \cdot \Sigma \cdot (x^1 - x^2)\]
D-efficiency: Balance Question Trade-off

- D-efficiency = Non-convex function $f(d, v)$ of distance: $d := \mu \cdot (x^1 - x^2)$
  variance: $v := (x^1 - x^2)' \cdot \Sigma \cdot (x^1 - x^2)$

Can evaluate $f(d, v)$ with 1-dim integral 😊

Piecewise Linear Interpolation

Optimal question selection = MIP
Computational Results for Question Selection

- Gaussian prior and 90% credibility ellipsoid, 100 inst.
  - 12 features, 2 profiles, 5 questions, 1-step Bayes

<table>
<thead>
<tr>
<th></th>
<th>Toubia et al.</th>
<th>PWL D-Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasible $\beta$</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>Distance (scaled)</td>
<td>0.97</td>
<td>0.85</td>
</tr>
<tr>
<td>D-Efficiency</td>
<td>2.2E+07</td>
<td>7.00E+06</td>
</tr>
<tr>
<td>Gaussian Volume</td>
<td>0.74</td>
<td>0.40</td>
</tr>
</tbody>
</table>

- 1 step for random covariance/ellipsoid

<table>
<thead>
<tr>
<th></th>
<th>Toubia et al.</th>
<th>PWL D-Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Efficiency</td>
<td>0.016</td>
<td>0.015</td>
</tr>
<tr>
<td>variance</td>
<td>110</td>
<td>83</td>
</tr>
<tr>
<td>distance</td>
<td>8.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Area R1 / R2</td>
<td>32% / 68%</td>
<td>47% / 53%</td>
</tr>
</tbody>
</table>
Summary

• **Messages:**
  – Always choose Chewbacca!
  – Polyhedral $\rightarrow$ Geometric $\approx$ Bayesian
    • Question selection and update with optimization and limited sampling (1-dim integrals)
    • Point estimation and credibility region
    • Improvements in point estimation, reduction of uncertainty and precision of credibility region
    • Also works for more profiles and attribute levels

• **Future:**
  – Combination and comparison with fully Bayesian
  – Combine with polyhedral updates
  – Computational improvements
  – Field experiments