Web-Based Adaptive Conjoint Analysis for Consumer Marketing, based on Polyhedral Geometry

with assistance from Olivier Toubia, Duncan Simester, and John Hauser

March 14, 2002
1. The Analytic Center of a Polyhedral System
2. Conjoint Analysis for Consumer Marketing
3. The Polyhedral Method
4. Computational Results
5. Conclusions
The Analytic Center of a Polyhedral System

Consider a polyhedral system of the form:

\[ Ax \leq b, \quad Mx = g, \]

The analytic center is the solution of:

\[
\text{(ACP:)} \quad \max_{x,s} \sum_{i=1}^{m} s_i \\
\text{s.t.} \quad Ax + s = b \\
\quad s \geq 0 \\
\quad Mx = g.
\]
(ACP:) \[ \max_{x,s} \prod_{i=1}^{m} s_i \]

s.t. \[ Ax + s = b \]
\[ s \geq 0 \]
\[ Mx = g . \]

is the same as:

(ACP:) \[ \min_{x,s} - \sum_{i=1}^{m} \ln(s_i) \]

s.t. \[ Ax + s = b \]
\[ s \geq 0 \]
\[ Mx = g . \]
The analytic center possesses a very nice “centrality” property.

Suppose that \((\hat{x}, \hat{s})\) is the analytic center.

Define:

\[
\hat{S}^{-2} := \begin{pmatrix}
(\hat{s}_1)^{-2} & 0 & \ldots & 0 \\
0 & (\hat{s}_2)^{-2} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & (\hat{s}_m)^{-2}
\end{pmatrix}.
\]
Define the following sets:

\[ P := \{ x \mid Mx = g, \ Ax \leq b \} \]

\[ E_{IN} := \left\{ x \mid Mx = g, \ (x - \hat{x})^T A^T \hat{S}^{-2} A(x - \hat{x}) \leq 1 \right\} \]

\[ E_{OUT} := \left\{ x \mid Mx = g, \ (x - \hat{x})^T A^T \hat{S}^{-2} A(x - \hat{x}) \leq m \right\} \]

**Theorem:** If \((\hat{x}, \hat{s})\) is the analytic center, then:

\[ E_{IN} \subset P \subset E_{OUT} . \]
Figure 1: Illustration of the Ellipsoid construction at the analytic center.
Theorem: If $(\hat{x}, \hat{s})$ is the analytic center, then:

\[ E_{\text{IN}} \subset P \subset E_{\text{OUT}}. \]

Note that $E_{\text{IN}}$ (or $E_{\text{OUT}}$) approximates the shape of $P$ to within a factor of $m$. 
Analytic Center

Centrality Property

...Bounding Ellipsoids

Figure 2: Ellipsoids.
Computation of Approximate Analytic Center

  - Newton’s method is very efficient in theory
- In practice, can solve in roughly 50% of the time it takes to solve an LP.
- Approximate analytic center has excellent properties as well.
Newton’s Method

\[(P:) \quad \min f(x)\]

\[x \in \mathbb{R}^n.\]

At \(x = \bar{x}\), \(f(x)\) can be approximated by:

\[f(x) \approx h(x) := f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^t H(\bar{x})(x - \bar{x}),\]

\(\nabla f(x)\) is the gradient of \(f(x)\)

\(H(x)\) is the Hessian of \(f(x)\).
Newton’s Method

(P:) \[ \min f(x) \approx h(x) := f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^T H(\bar{x})(x - \bar{x}) \]

\[ x \in \mathbb{R}^n. \]

Let us solve \( \nabla h(x) = 0. \)

\[ \nabla h(x) = \nabla f(\bar{x}) + H(\bar{x})(x - \bar{x}) \]
Newton’s Method

We therefore solve:

\[ \nabla f(\bar{x}) + H(\bar{x})(x - \bar{x}) = 0 \]

\[ d := x - \bar{x} = -H(\bar{x})^{-1}\nabla f(\bar{x}). \]

\(-H(\bar{x})^{-1}\nabla f(\bar{x})\) is called the Newton direction or the Newton step at \( x = \bar{x} \).
Algorithm for \((P:)
\)

Newton’s Method:

**Step 0** Given \(x^0\), set \(k \leftarrow 0\)

**Step 1** \(d^k = -H(x^k)^{-1} \nabla f(x^k)\). If \(d^k = 0\), then stop.

**Step 2** Choose stepsize \(\alpha^k = 1\).

**Step 3** Set \(x^{k+1} \leftarrow x^k + \alpha^k d^k\), \(k \leftarrow k + 1\). Go to **Step 1**.
Newton’s Method

Proposition 0.0.1 If $H(x)$ is SPD, then $d = -H(x)^{-1}\nabla f(x)$ is a descent direction, i.e. $f(x + \alpha d) < f(x)$ for all sufficiently small values of $\alpha$.

- The method assumes $H(x^k)$ is nonsingular at each iteration.
- There is no guarantee that $f(x^{k+1}) \leq f(x^k)$.
- Step 2 could be augmented by a linesearch of $f(x^k + \alpha d^k)$ to find an optimal value of the stepsize parameter $\alpha$. 
Web-Based Adaptive Conjoint Analysis

by Olivier Toubia
Duncan Simester
John Hauser
http://mitsloan.mit.edu/vc

Latest Draft: February 26, 2002
• Motivation – feature selection for product design
• Products are more complex
• Products have shorter life cycles
• Web-based consumer surveys are cheap and “easy”
• Current Standard:
  – Sawtooth ACA (from Sawtooth Software, Inc.)
Product has $n$ binary features (think $n = 20$), such as:

- price (low, high)
- size (small, big)
- weight (light, heavy)
- outside key pocket (yes/no)
Each consumer has an additive nonnegative utility for each feature:

Eric: \((u_1, u_2, \ldots, u_n) = (1, 7, 23, \ldots, 7.93, 14.2)\)

May Ling: \((u_1, u_2, \ldots, u_n) = (5, 9, 28, \ldots, 1.14, 54.2)\)

If product has features 1, 7, 8, 10, 11, 15, 19, utility is:

\[u_1 + u_7 + u_8 + u_{10} + u_{11} + u_{15} + u_{19}.\]
We would like to accurately estimate

\[ u = (u_1, \ldots, u_n) \]

for each consumer in our sample.

How to do this?
### Paired Comparison Questions

<table>
<thead>
<tr>
<th>Feature</th>
<th>Product Version A</th>
<th>Product Version B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Please rank $A$ vs. $B$:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>-4</td>
</tr>
<tr>
<td>-4</td>
<td>-3</td>
</tr>
<tr>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td>-2</td>
<td>-1</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Eric
Paired Comparison Questions

<table>
<thead>
<tr>
<th>Feature</th>
<th>Product Version A</th>
<th>Product Version B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Please rank $A$ vs. $B$:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>5</td>
</tr>
<tr>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$b^T \bar{u} - a^T \bar{u} = -4$$

$$(b - a)^T \bar{u} = -4$$

$$(1, 0, -1, 1, \ldots, 0 \text{ or } 1 \text{ or } -1)^T \bar{u} = g$$
Suppose we have asked and received answers to five questions.

\( \bar{u} = (\bar{u}_1, \ldots, \bar{u}_n) \) must satisfy:

\[
\begin{align*}
  m_1^T \bar{u} & = g_1 \\
  m_2^T \bar{u} & = g_2 \\
  & \vdots \\
  m_5^T \bar{u} & = g_5 \\
  \bar{u} & \geq 0.
\end{align*}
\]

\( \bar{u} \in P_5 = \{ u | m_1^T u = g_1, \ldots, m_5^T u = g_5, u \geq 0 \} \)

\( P_5 \) is a polyhedron of dimension \( n - 5 \).
\( \vec{u} \in P_5 = \{ u | m_1^T u = g_1, \ldots, m_5^T u = g_5, u \geq 0 \} \)

Now I ask a sixth question and receive an answer.

\[ m_6^T \vec{u} = g_6 \]

Now I know \( \vec{u} \in P_6 = \{ u | m_1^T u = g_1, \ldots, m_6^T u = g_6, u \geq 0 \} \)
We would like to choose the sixth question in such a way that we minimize the volume (the area in 2-dimensions) of $P_6$.

This is not so easy, but not so hard either...
Polyhedral Geometry

...Minimization of Volume...
...Minimization of Volume...
Better question

Minimization of Volume
• Compute and enumerate all extreme points of $P_5$
  \[ \nu^1, \ldots, \nu^{857} \]
• Find the longest “axis” of $P_5$
  \[ \text{maximize}_{i,j} \ |\nu^i - \nu^j| \]
• Compute and enumerate all extreme points of $P_5$
  $\mathbf{v}^1, \ldots, \mathbf{v}^{857}$

• Find the longest “axis” of $P_5$

$$\max_{i,j} ||\mathbf{v}^i - \mathbf{v}^j|| = ||\mathbf{v}^{193} - \mathbf{v}^{574}||$$
• Choose the next question $m_6$ so that its hyperplane is most perpendicular to $v_1^{193} - v_1^{574}$

$$\text{maximize}_{m_i} \frac{|m_i^T (v_1^{193} - v_1^{574})|}{||m_i||} \text{ over all questions } m_i = b_i - a_i$$
• This works well with \( n \leq 9 \) features

• Number of extreme points grows exponentially in \( n \)

• Computation of \( m_6 \) must be done in at most two seconds.
Current Strategy: Analytic Center

- Compute analytic center $\hat{x}$ of $P$

- Construct circumscribed ellipsoid $E_{OUT}$
  
  - $E_{OUT}$ is presumably a good approximator of the shape of $P$

- Compute longest axis of $E_{OUT}$: $\hat{\nu}$
Current Strategy: Analytic Center
Current Strategy: Analytic Center
• Now choose next question \( m_6 \) to be most perpendicular to the axis \( \hat{v} \).

\[
\text{maximize}_{m_i} \frac{|m_i^T \hat{v}|}{||m_i||}
\]

• Even with \( n = 30 \) features, computation takes less than 1/2 second on a modest PC.
Illustration of Geometry
ACA

Illustration of Geometry
• Inaccurate responses
• Inconsistent responses
• Non-binary discrete features
• Restrictions on types of questions
• First question – priors
• Programming – MATLAB
Illustration of Geometry

Website Demo

http://mitsloan.mit.edu/vc
Conclusions

- Analytic Center strategy is very promising

- Value of advanced ideas in optimization theory