Ellipsoidal Methods for Adaptive Choice-based Conjoint Analysis (CBCA)

Juan Pablo Vielma

Massachusetts Institute of Technology

Joint work with Denis Saure

Operations Management Seminar,
Rotman School of Management,
Toronto, Canada. December, 2016.

Motivation: (Custom) Product Recommendations

Feature	SX530	RX100
Zoom	50x	3.6x
Prize	$249.99	$399.99
Weight	15.68 ounces	7.5 ounces
Prefer	☑	☐

Feature	TG-4	G9
Waterproof	Yes	No
Prize	$249.99	$399.99
Weight	7.36 lb	7.5 lb
Prefer	☐	☑

Feature	TG-4	Galaxy 2
Waterproof	Yes	No
Prize	$249.99	$399.99
Viewfinder	Electronic	Optical
Prefer	☑	☐

We recommend:

Towards CBCA-Based Recommendations

- Individual preference estimates with few questions
- Adaptive Questions:
  - Fast question selection
  - Pick **next** question to reduce uncertainty
  - Quantify estimate variance
- Favorable properties for future:
  - Intuitive geometric model (e.g. Robust Opt.)
  - Parametric model

Choice-based Conjoint Analysis

Feature	Chewbacca	BB-8
Wookiee	Yes	No
Droid	No	Yes
Blaster	Yes	No
I would buy toy	☑	☐
Product Profile	x^1	x^2

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x^2$$

MNL Preference Model

- Utilities for 2 products, d features

$$U_1 = \beta \cdot x^1 + \epsilon_1 = \sum_{i=1}^{d} \beta_i x_i^1 + \epsilon_1$$

$$U_2 = \beta \cdot x^2 + \epsilon_2 = \sum_{i=1}^{d} \beta_i x_i^2 + \epsilon_2$$

part-worths

product profile

noise (gumbel)

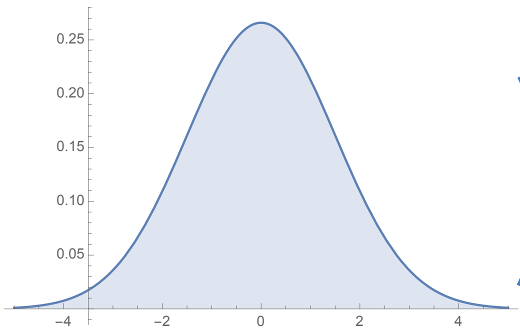
- Utility maximizing customer: $x^1 \succeq x^2 \Leftrightarrow U_1 \text{“}\geq\text{”} U_2$

- 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

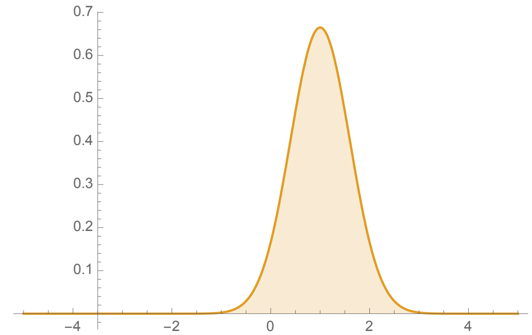
Prior Distribution of β



Feature	SX530	RX100
Zoom	50x	3.6x
Prize	$249.99	$399.99
Weight	15.68 ounces	7.5 ounces
Prefer	☑	☐

Bayesian Update

Posterior Distribution

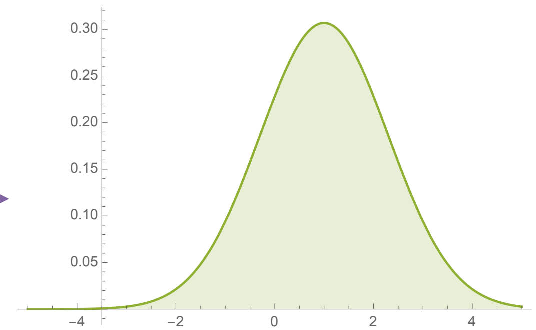


MCMC

Feature	TG-4	Galaxy 2
Waterproof	Yes	No
Prize	$249.99	$399.99
Viewfinder	Electronic	Optical
Prefer	☑	☐

Bayesian Update

Posterior Distribution

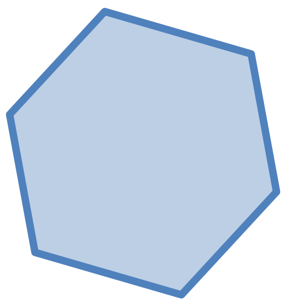


- Update uses MNL response error ✓
- Question Selection (even knowing answer): Enumeration ✗

Next Question To Reduce “Variance”: Polyhedral

Toubia, Hauser and Simester, ‘04

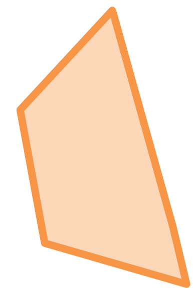
Polyhedron Containing β



Feature	SX530	RX100
Zoom	50x	3.6x
Prize	$249.99	$399.99
Weight	15.68 ounces	7.5 ounces
Prefer	✓	☐

Geometric Update

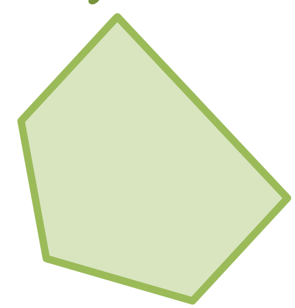
Posterior Polyhedron



Feature	TG-4	Galaxy 2
Waterproof	Yes	No
Prize	$249.99	$399.99
Viewfinder	Electronic	Optical
Prefer	✓	☐

Geometric Update

Posterior Polyhedron



- Update ignores response error ✗
- Question Selection: (Multi-Obj.) Discrete Optimization ✓

Outline

- Objective: Combine Bayesian and polyhedral methods into intuitive geometric approach

1. Review of geometry of polyhedral method
2. Incorporating response error = ellipsoids
3. MIP based near-optimal question selection to reduce variance measure (D-efficiency)
4. Optimal one-step look-ahead moment-matching approximate Bayesian approach (~~OOLMMABA?~~)

Ellipsoidal Method

Polyhedral Method

Preference Model and Geometric Interpretation

- Utilities for 2 products, d features, logit model

$$U_1 = \beta \cdot x^1 + \epsilon_1 = \sum_{i=1}^{d} \beta_i x_i^1 + \epsilon_1$$

$$U_2 = \beta \cdot x^2 + \epsilon_2 = \sum_{i=1}^{d} \beta_i x_i^2 + \epsilon_2$$

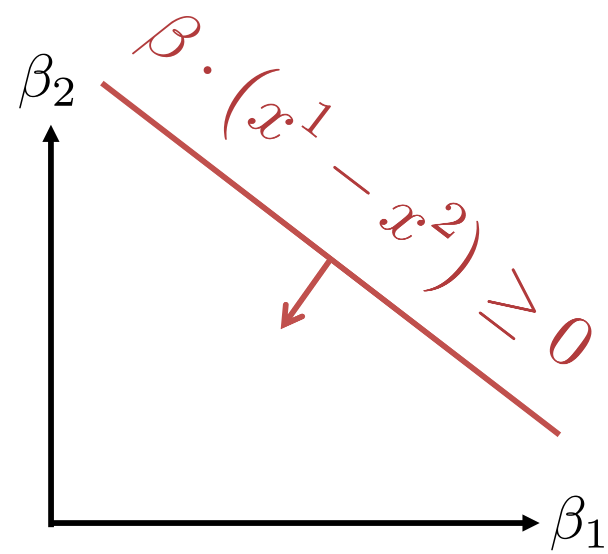
part-worths

product profile

noise (gumbel)

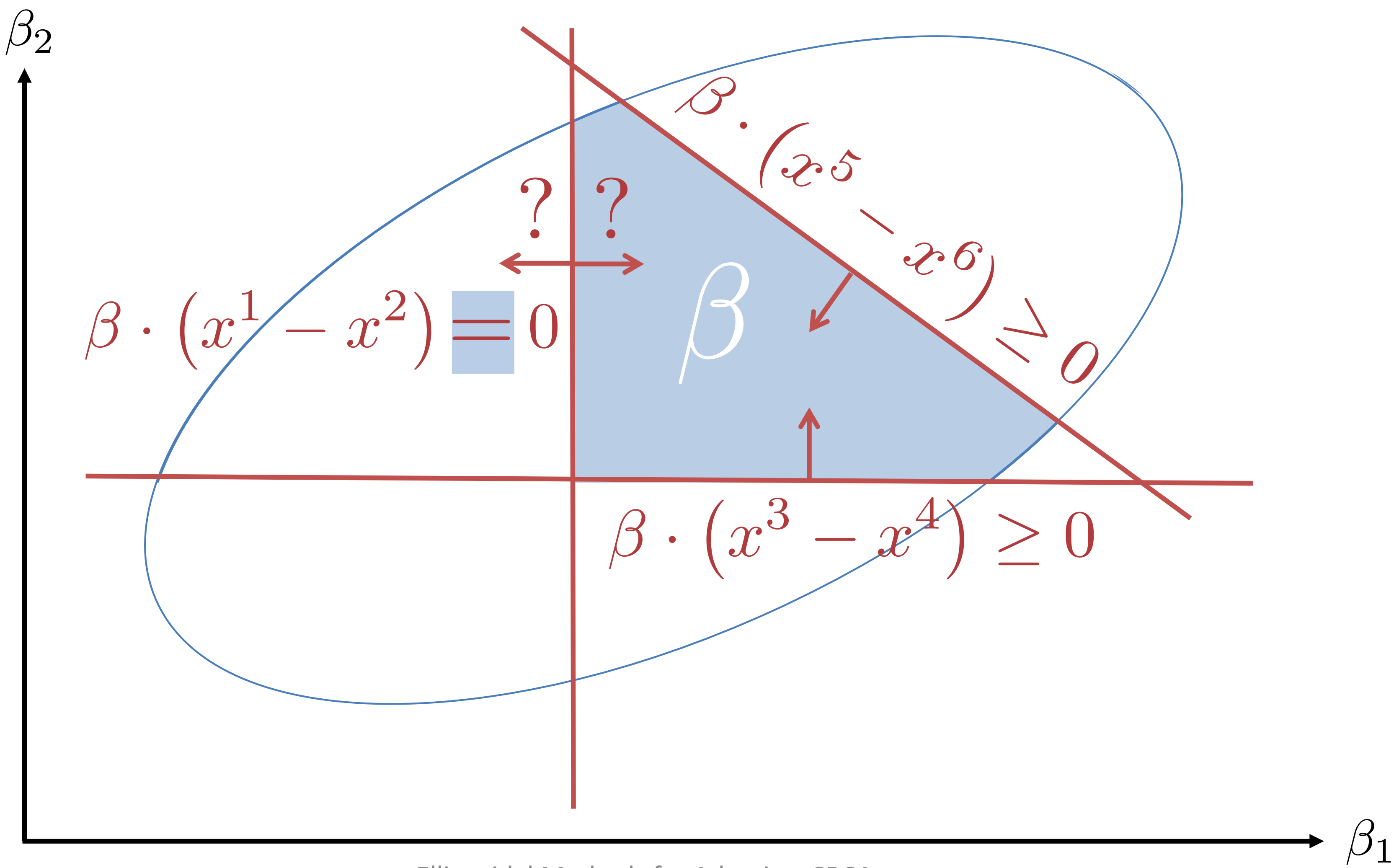
- Utility maximizing customer
  - Geometric interpretation of preference for product 1 without error

$$x^1 \succeq x^2 \Leftrightarrow U_1 \geq U_2$$



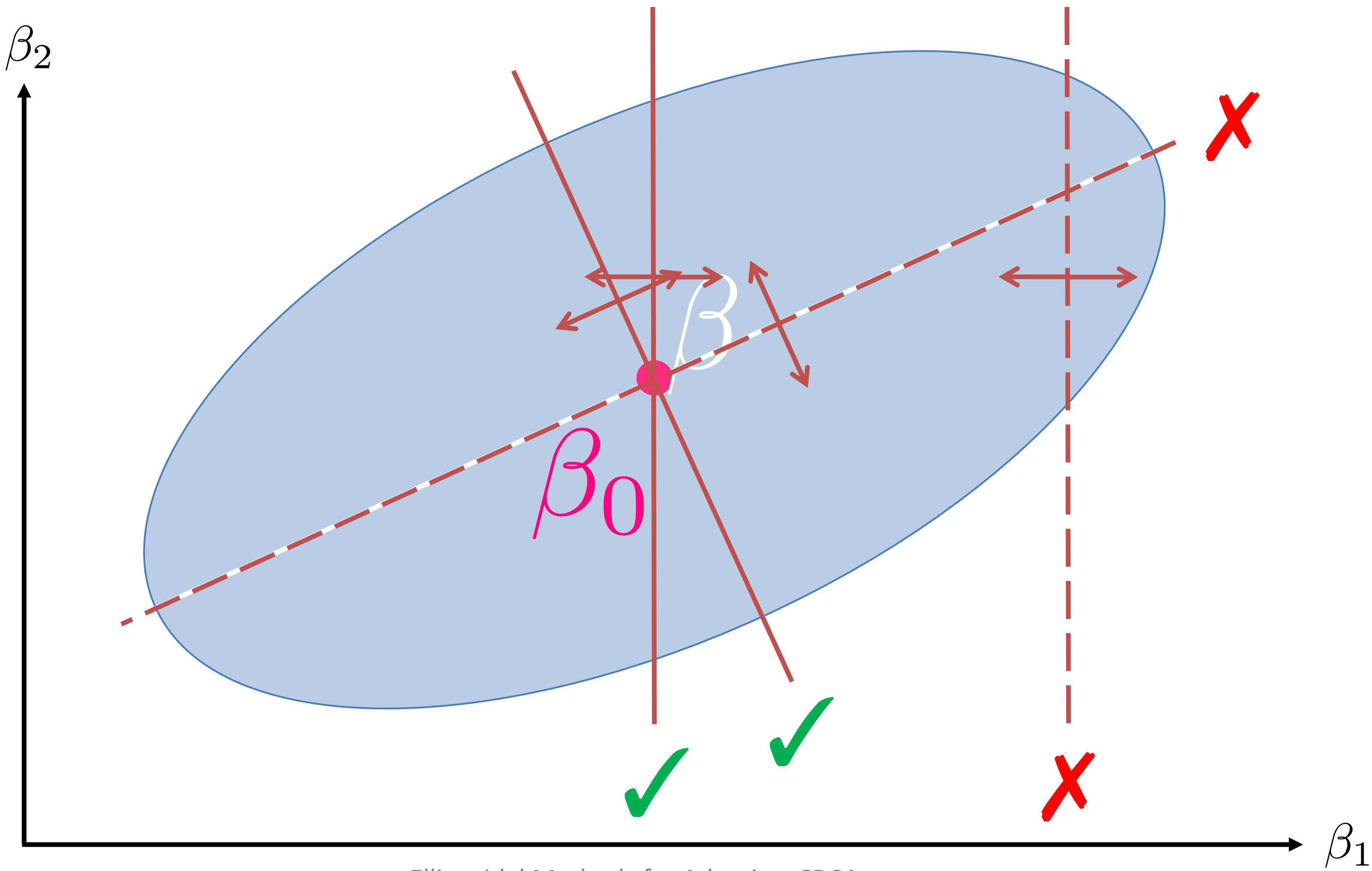
Polyhedral Method: Ask Question and Update

Geometric prior for β → $x^1 \succeq x^2$ → 2nd geometric posterior for β / 3rd geometric posterior for β



Polyhedral: Estimation and Question Selection

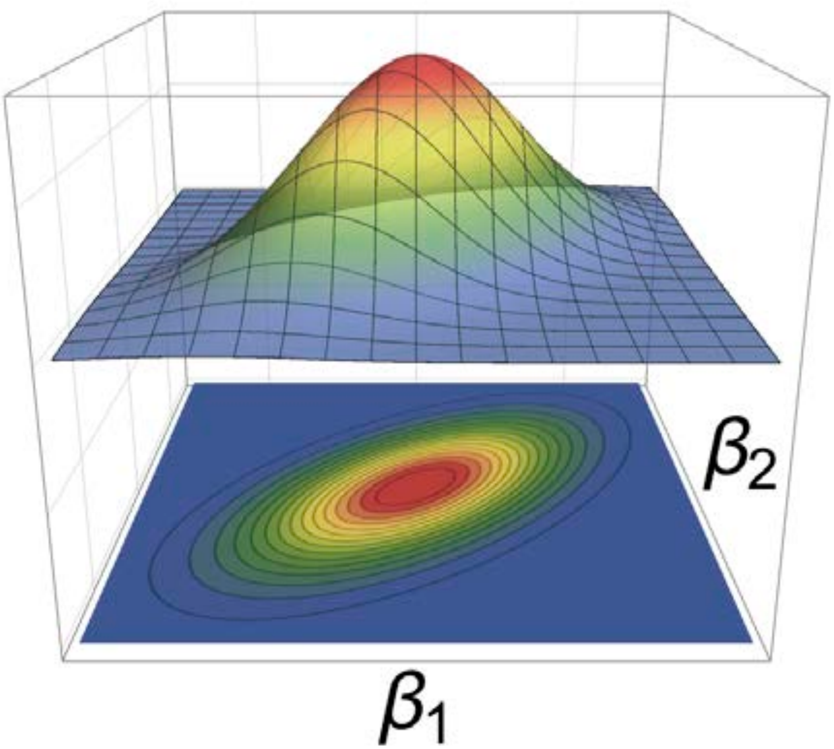
Good Question? estimator for β? 1. 2. Centroid ellipsoid ... symmetry β_0



Incorporating Response Error

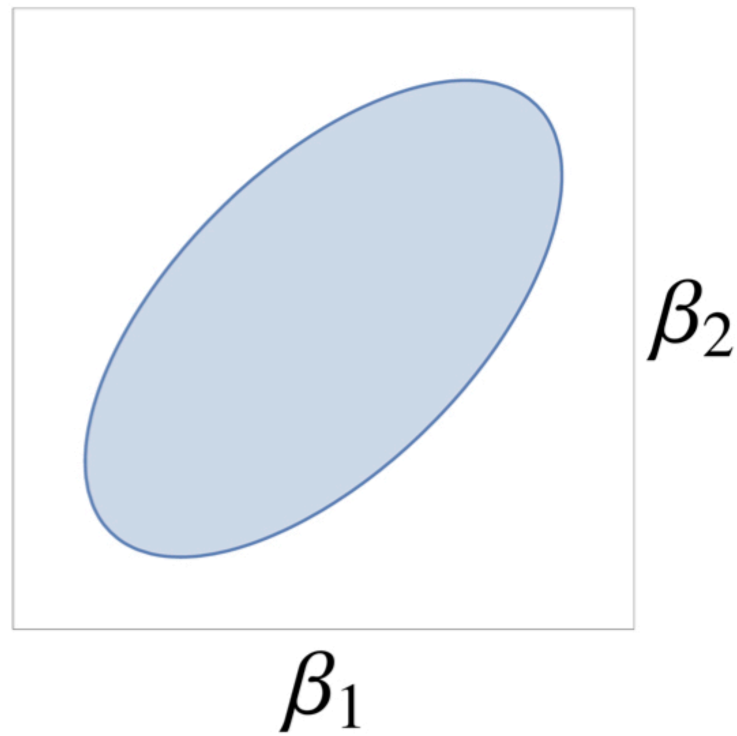
Distributions and Credibility Ellipsoids

Prior distribution of β



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

90% confidence/credibility ellipsoid

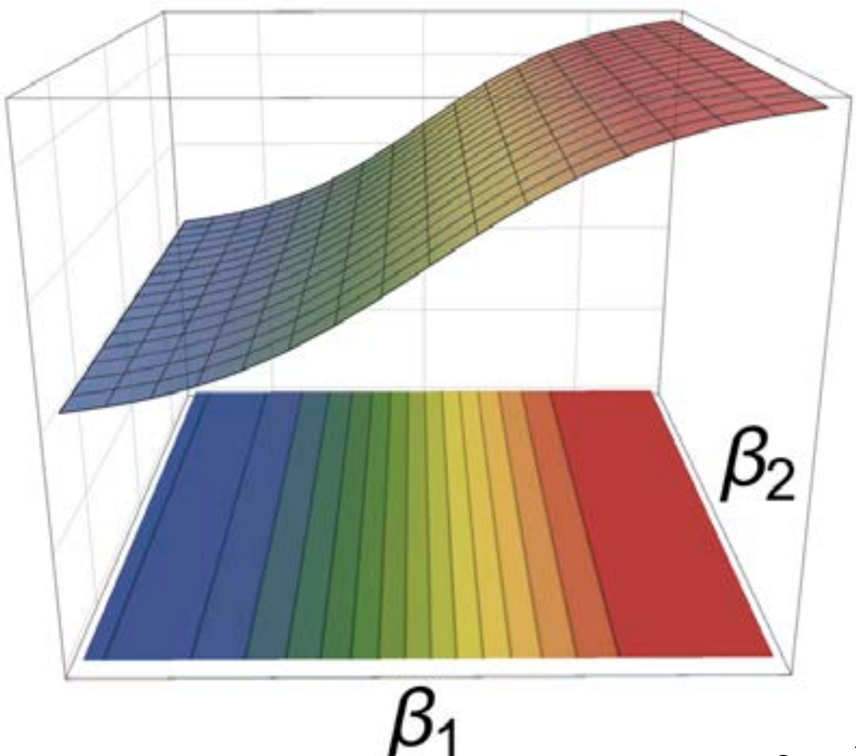


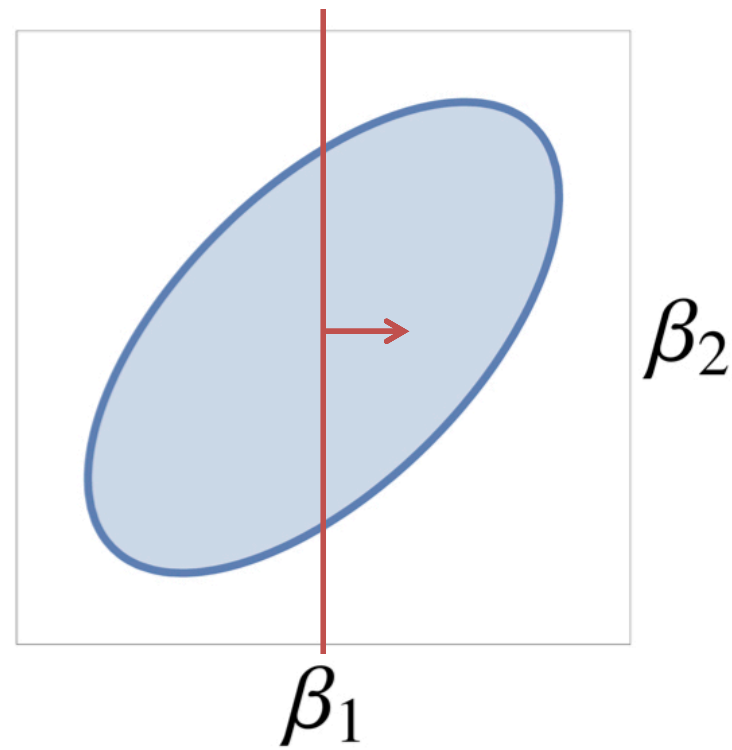
$$(\beta - \mu)' \cdot \Sigma^{-1} \cdot (\beta - \mu) \leq r$$

Answers with Error: Logit Probabilities

Likelihood Function

Question/Answer

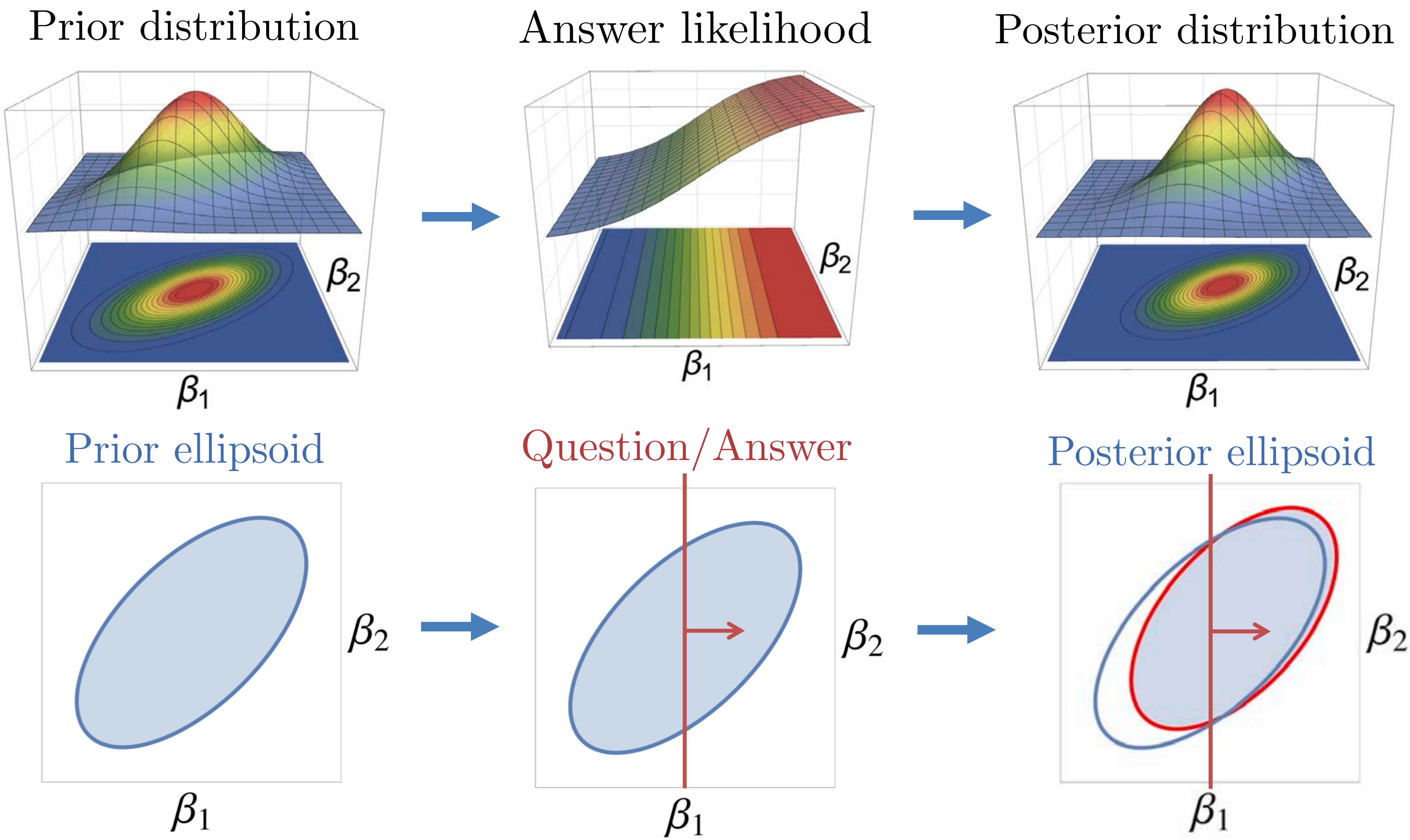




$$\mathbb{P}\left(x^1 \succeq x^2 \mid \beta\right) = \frac{e^{\beta \cdot x^1}}{e^{\beta \cdot x^1} + e^{\beta \cdot x^2}}$$

$$x^1 \succeq x^2$$

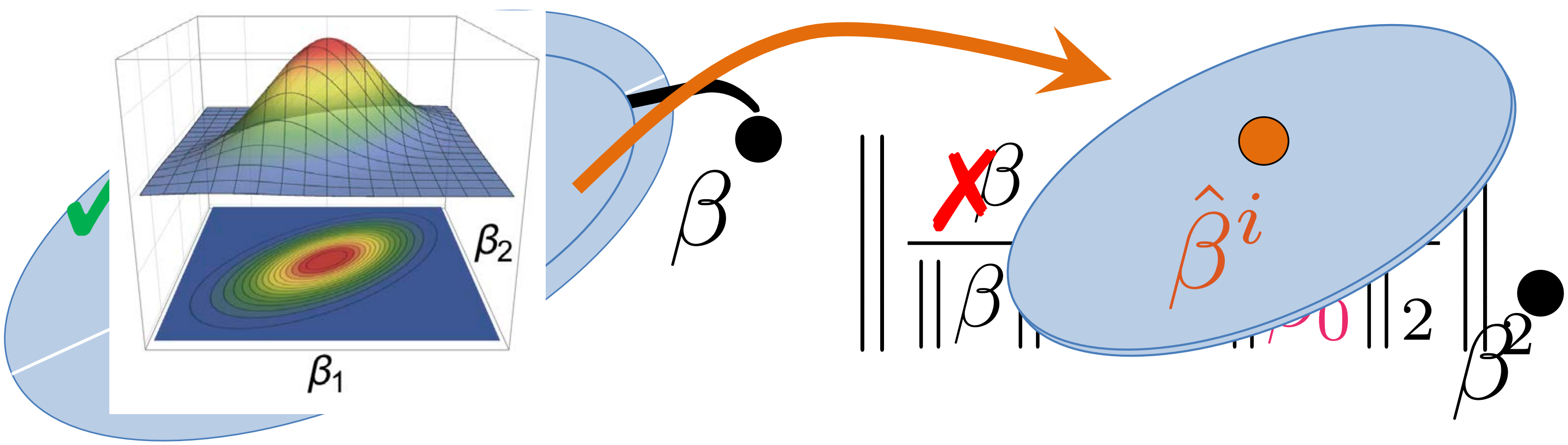
Bayesian Update and Geometric Updates



Computational Comparison of Updates

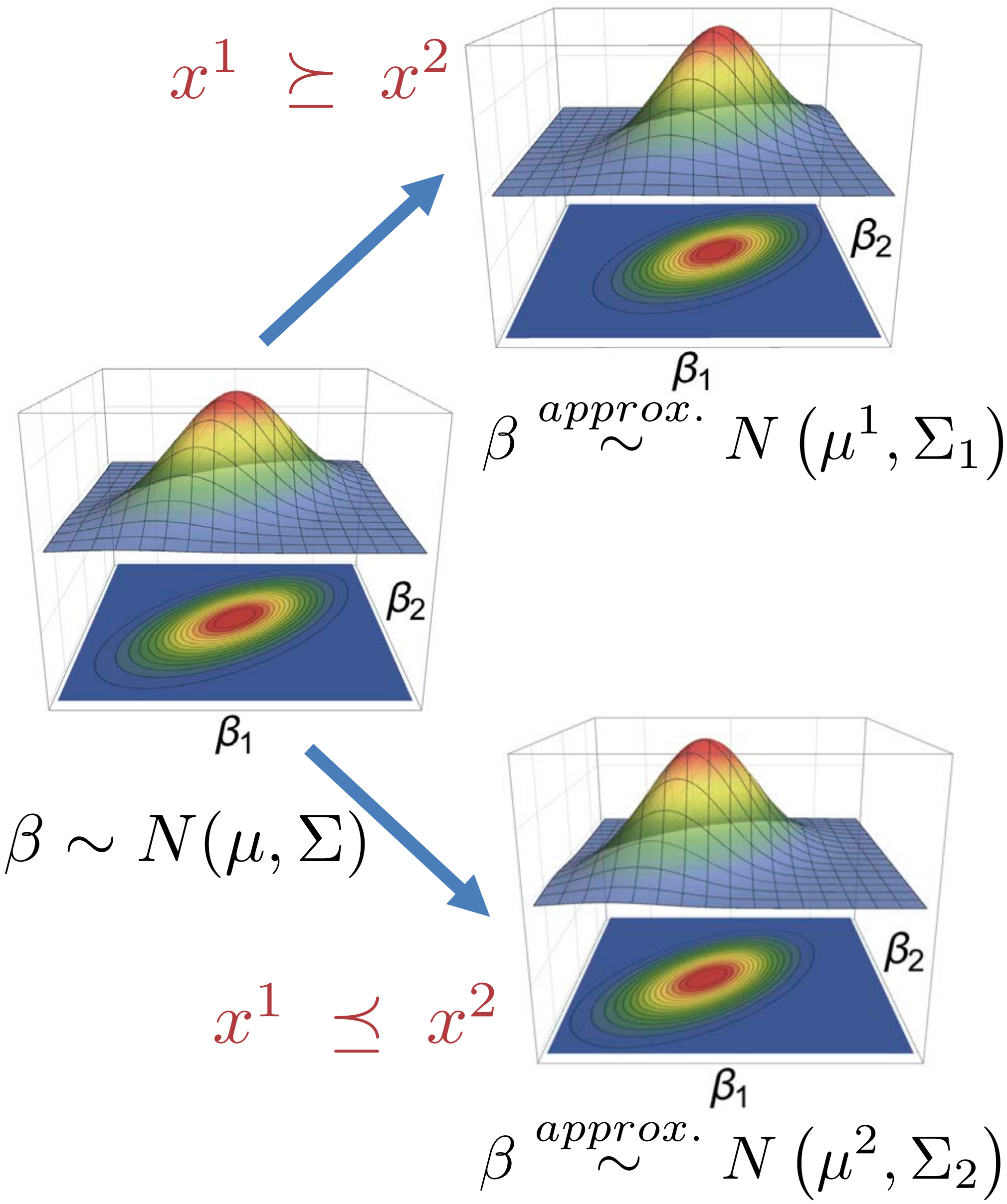
- Gaussian prior and 90% credibility ellipsoid, 100 inst.
  - 12 features, 2 profiles and 5 questions
  - Sample 1 “true” β and simulate MNL responses with it

	Polyhedral	Ellipsoidal
Feasible β	0.53	0.93
Distance (scaled)	0.92	0.85
Gaussian Volume	0.03	0.40

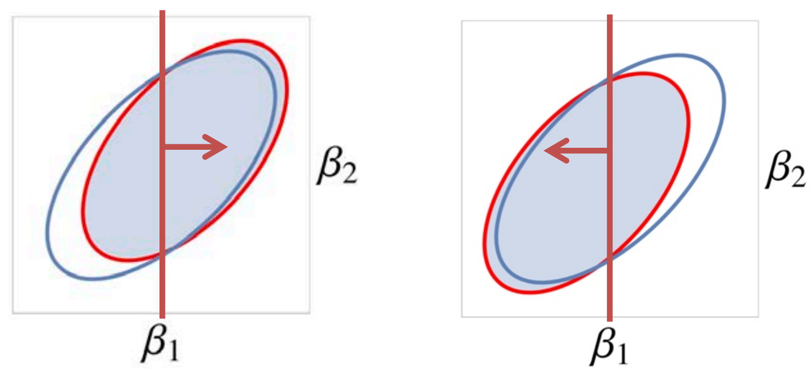


Question Selection: Optimizing D-Efficiency

D-Efficiency and Posterior Covariance Matrix



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



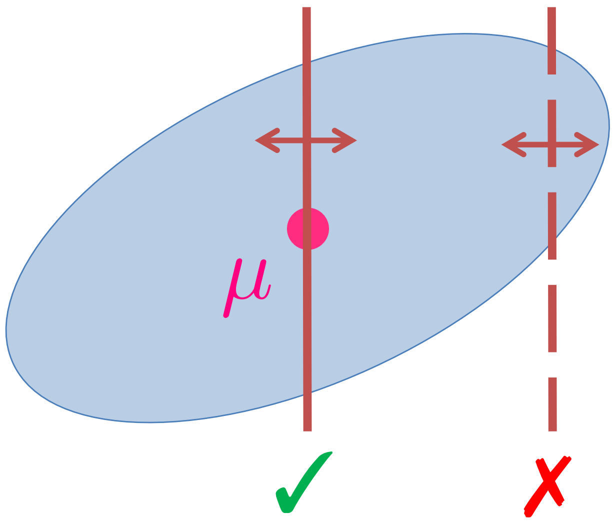
- Evaluating = multi-dim integration

Back to Question Selection: Property Trade-off

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

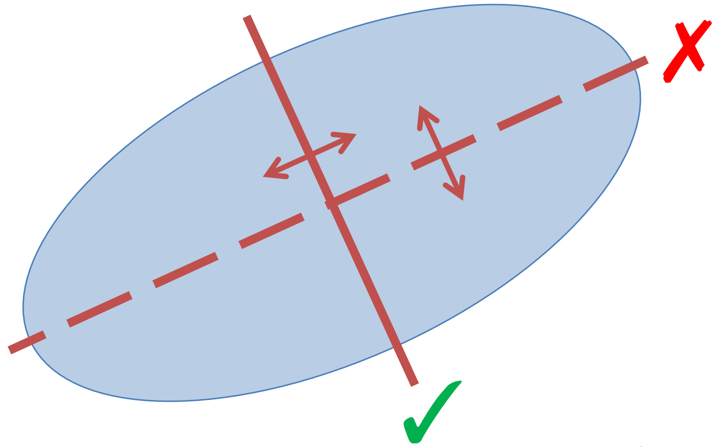
- Choice balance:
  - Minimize distance to center

$$\mu \cdot (x^1 - x^2)$$



- Postchoice symmetry:
  - Maximize variance of question

$$(x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2)$$

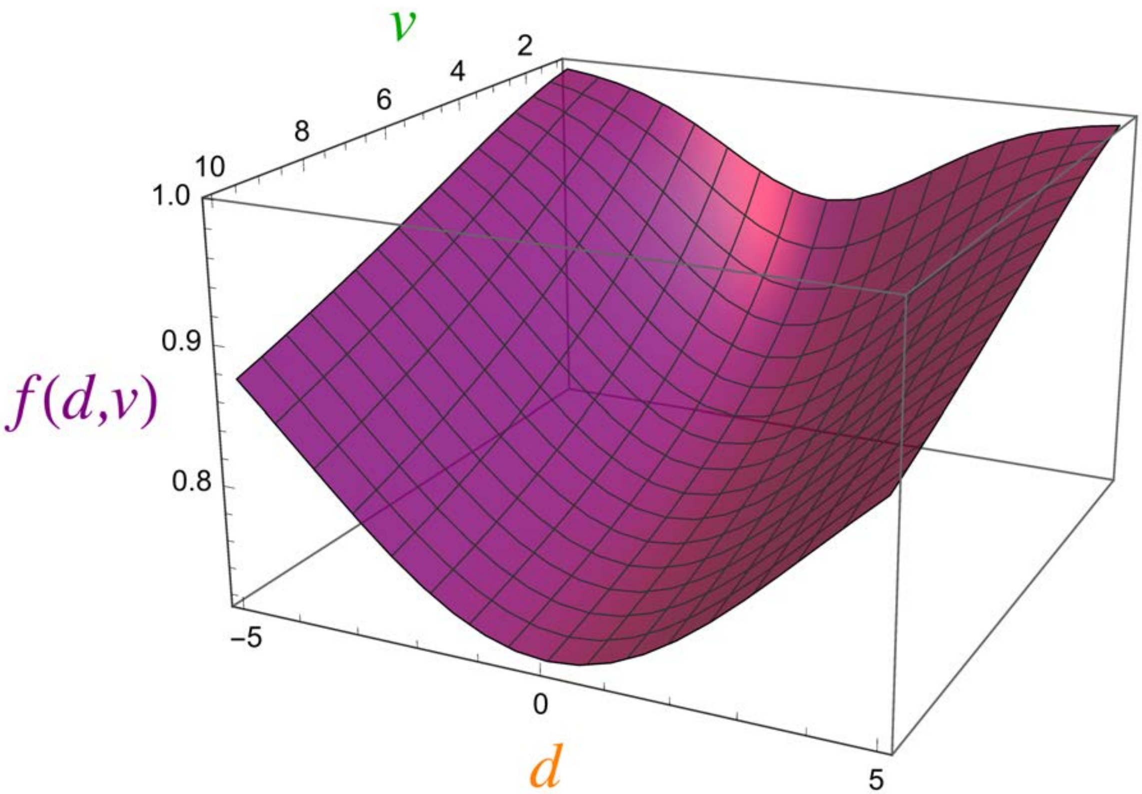


D-efficiency Simplification for CBCA

- D-efficiency = Non-convex function $f(d, v)$ of

  distance: $d := \mu \cdot (x^1 - x^2)$

  variance: $v := (x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2)$



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

Optimization Model

$$\min \quad f(d, v) \quad ✗$$

$$s.t.$$

$$\mu \cdot \left(x^1 - x^2\right) = d \quad ✓$$

$$\left(x^1 - x^2\right)' \cdot \Sigma \cdot \left(x^1 - x^2\right) = v \quad ✗✓$$

$$A^1 x^1 + A^2 x^2 \leq b \quad ✓$$

$$x^1 \neq x^2 \quad ✗✓$$

linearize $x_i^k \cdot x_j^l$

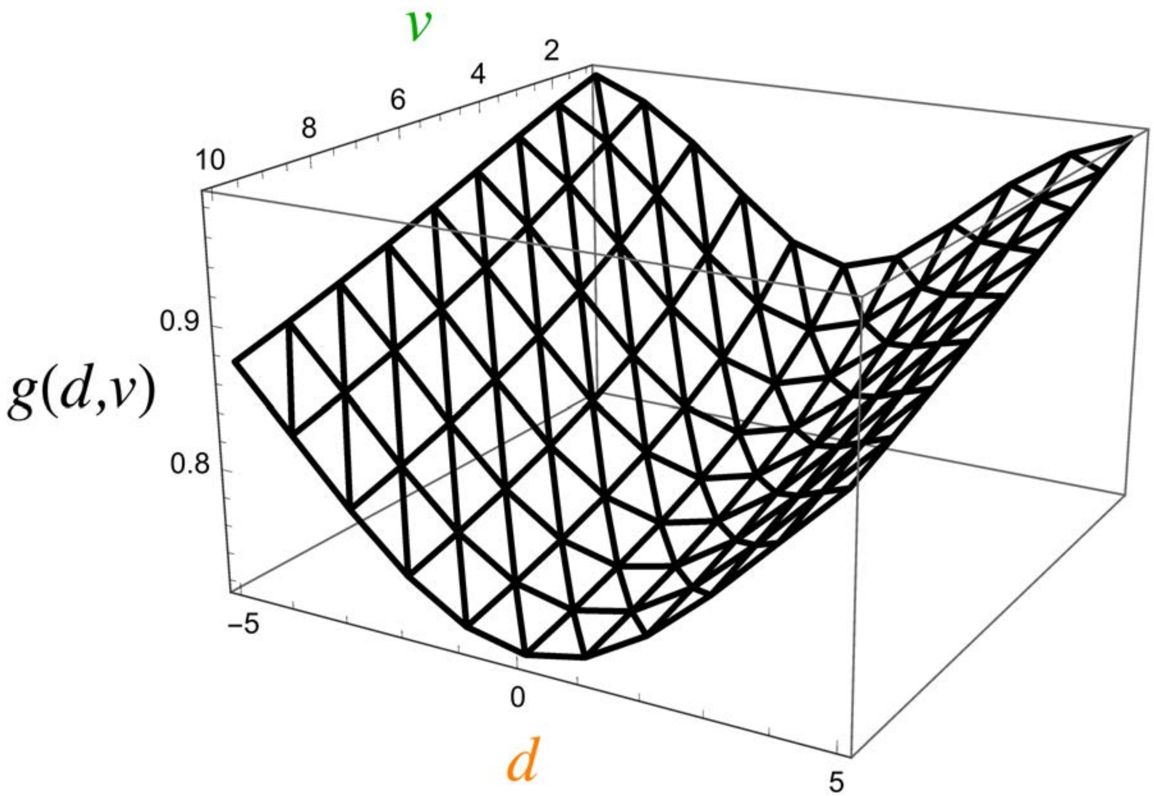
$$x^1, x^2 \in \{0, 1\}^n$$

Piecewise Linear Approximation

- D-efficiency = Non-convex function $f(d, v)$ of

  distance: $d := \mu \cdot (x^1 - x^2)$

  variance: $v := (x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2)$



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

Piecewise Linear Interpolation

MIP formulation

Putting Everything Together: Ellipsoidal Method

MIP-based Adaptive Questionnaires

Feature	SX530	RX100
Zoom	50x	3.6x
Prize	$249.99	$399.99
Weight	15.68 ounces	7.5 ounces
Prefer	☑	☐

Feature	TG-4	G9
Waterproof	Yes	No
Prize	$249.99	$399.99
Weight	7.36 lb	7.5 lb
Prefer	☐	☑

Feature	TG-4	Galaxy 2
Waterproof	Yes	No
Prize	$249.99	$399.99
Viewfinder	Electronic	Optical
Prefer	☑	☐

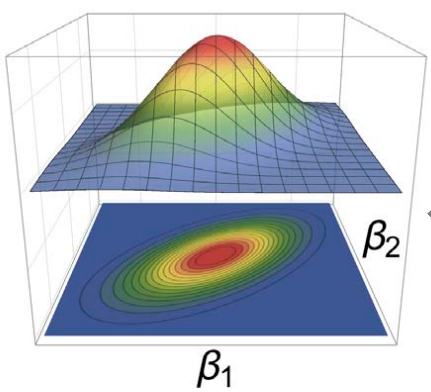
$$\mathbb{E}\left(\beta \mid Y, X^1, X^2\right)$$

$$\operatorname{cov}\left(\beta \mid Y, X^1, X^2\right)$$

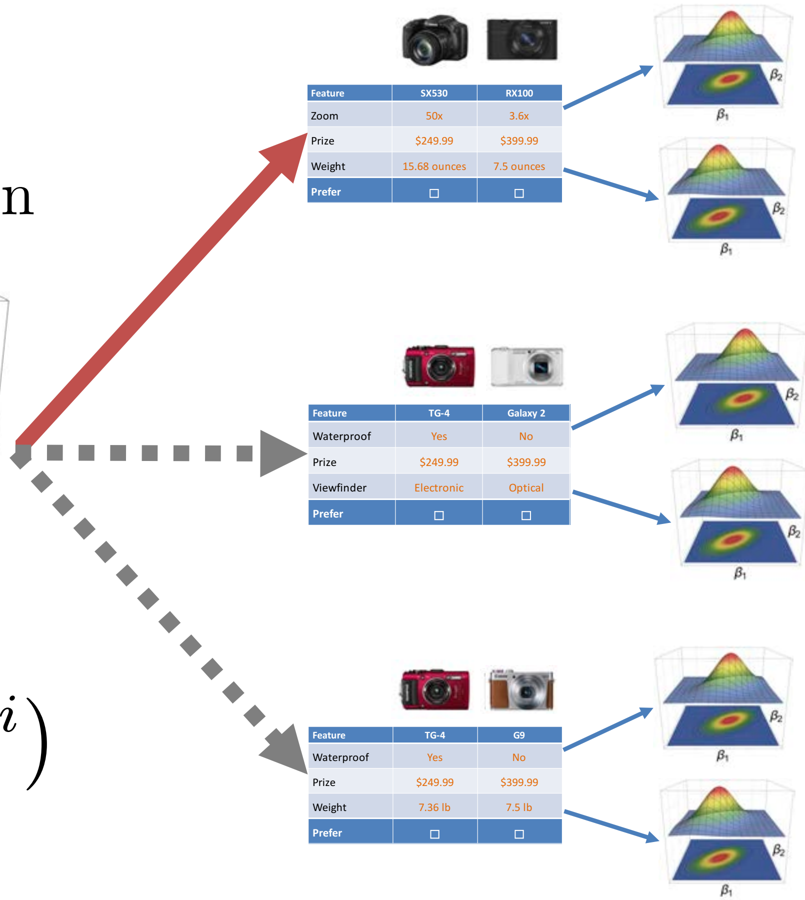
- Optimal one-step look-ahead moment-matching approximate Bayesian approach.

Optimal One-Step Look-Ahead

Prior distribution



$$\beta \sim N\left(\mu^i, \Sigma^i\right)$$



$$\min_{x^1, x^2} f\left(x^1, x^2\right)$$

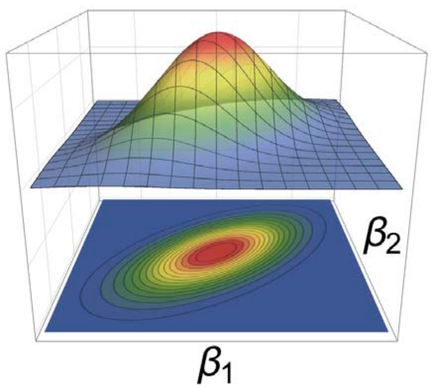
- Solve with MIP formulation

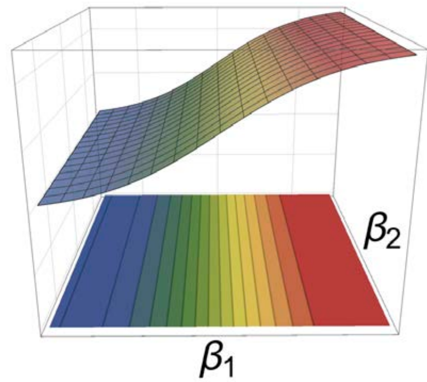
Moment-Matching Approximate Bayesian Update

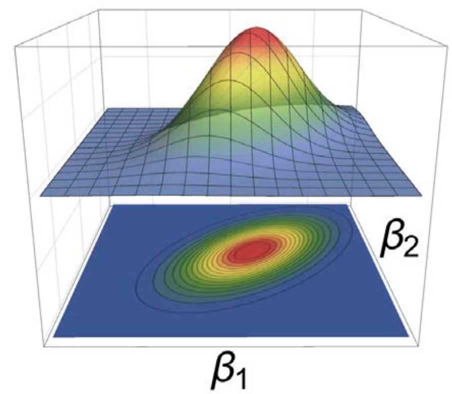
Answer likelihood

Prior distribution

Posterior distribution







Feature	TG-4	Galaxy 2
Waterproof	Yes	No
Prize	$249.99	$399.99
Viewfinder	Electron	Optical
Prefer	✓	□

$$\beta \sim N\left(\mu^i, \Sigma^i\right)$$

$$\beta \overset{approx.}{\sim} N\left(\mu^{i+1}, \Sigma^{i+1}\right)$$

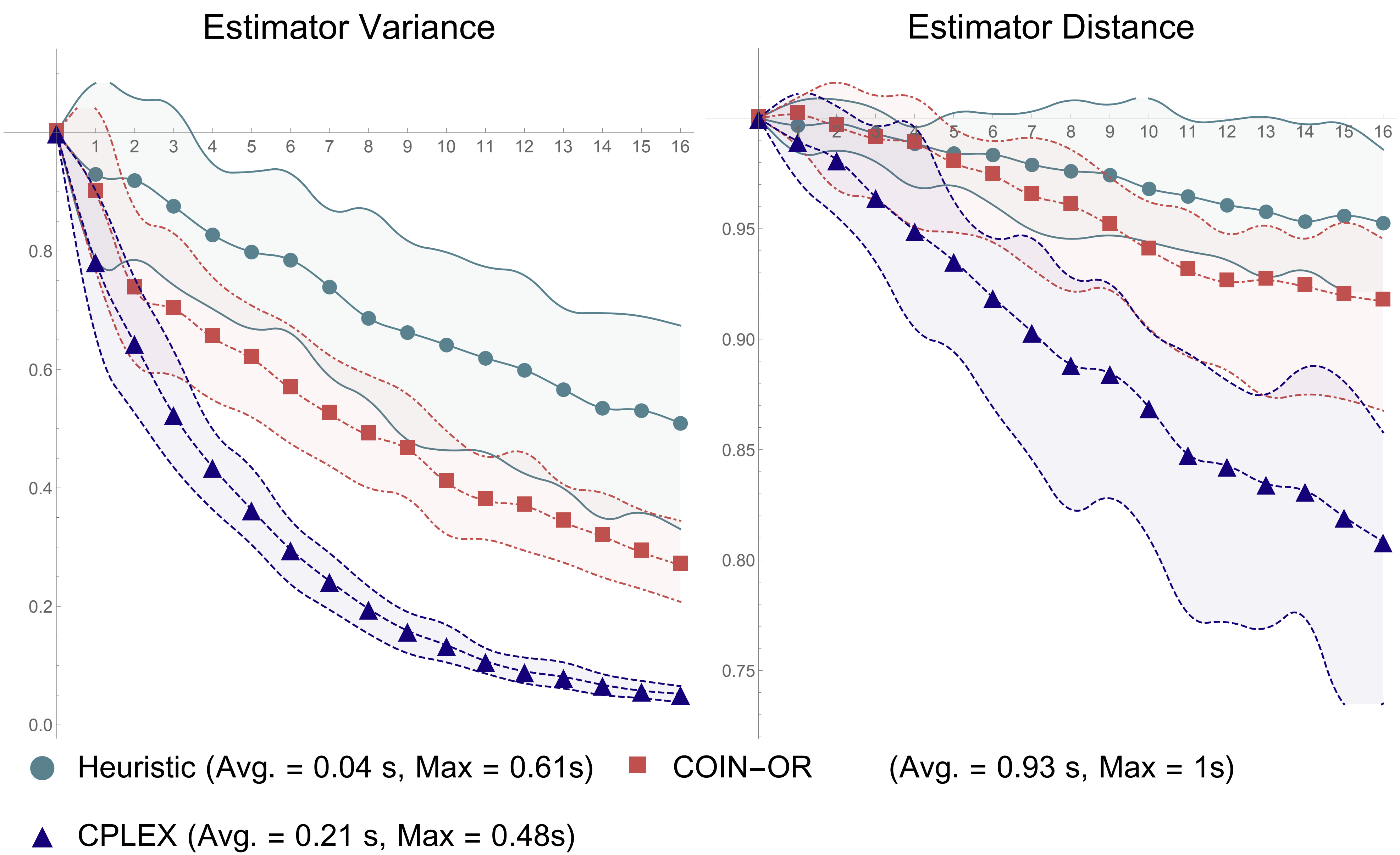
- $\mu^{i+1} = \mathbb{E}\left(\beta \,|\, y, x^1, x^2\right)$
- $\Sigma^{i+1} = \operatorname{cov}\left(\beta \,|\, y, x^1, x^2\right)$

1-dim integral

Computational Experiments

- 16 questions, 2 options, 12 features
- Simulate MNL responses with known β^*
- Question Selection
  - MIP-based using **CPLEX** and open source **COIN-OR** solver
  - Knapsack-based geometric **Heuristic** by Toubia et al.
- Time limits of 1 s
- Metrics:
  - Estimator variance = $\left(\det\operatorname{cov}\left(\beta \,|\, Y, X^1, X^2\right)\right)^{1/2}$
  - Estimator distance = $\left\|\mathbb{E}\left(\beta \,|\, Y, X^1, X^2\right) - \beta^*\right\|_2$
  - Computed for true posterior with MCMC

Results for 12 Features, 1 s time limit



Summary

- Messages:
  - Always choose Chewbacca!
  - Polyhedral → Geometric ≈ 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, attribute levels, etc.
- Future:
  - Combination and comparison with fully Bayesian
  - Combine with polyhedral updates
  - Computational improvements
  - Field experiments