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

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:

(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

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:

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

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

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

$$U_2 = \underbrace{\beta}_{\text{part-worths}} \cdot \underbrace{x^2}_{\text{product profile}} + \underbrace{\epsilon_2}_{\text{noise (gumbel)}} = \sum_{i=1}^{d} \beta_i x_i^2 + \epsilon_2$$

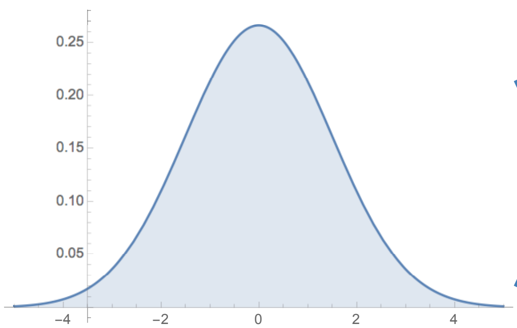
- Utility maximizing customer: $x^1 \succeq x^2 \Leftrightarrow U_1 \text{“}\geq\text{”} U_2$

- 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

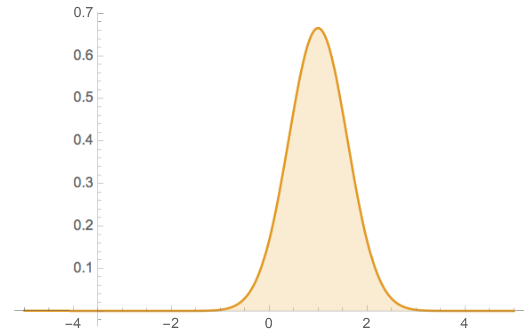
Prior Distribution of β



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

Bayesian Update

Posterior Distribution

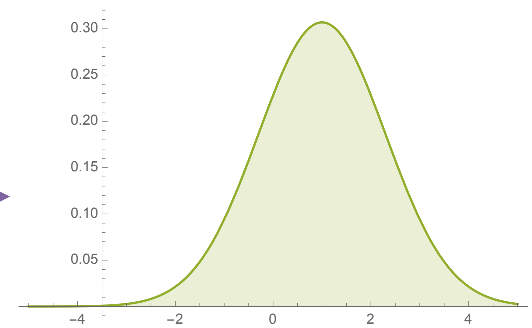


MCMC

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

Bayesian Update

Posterior Distribution

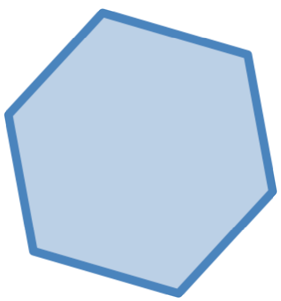


- Update uses MNL response error ✓
- Question Selection: Enumeration ✗
- Recommendation: Risk-averse Stochastic Optimization ✗

Next Question To Reduce "Variance": Polyhedral

Toubia, Hauser and Simester, '04

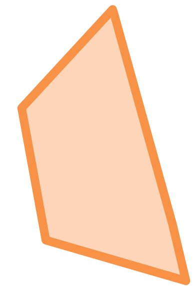
Polyhedron Containing β



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

Geometric Update

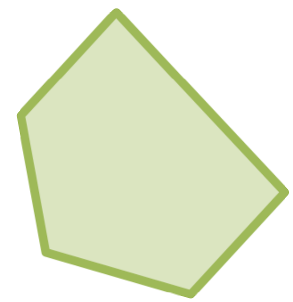
Posterior Polyhedron



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 ✓
- Recommendation: Robust Optimization ✓

“Improving” the 2004 Polyhedral Method

- More “re-interpreting” ideas from Toubia, Hauser and Simester, ‘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

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

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

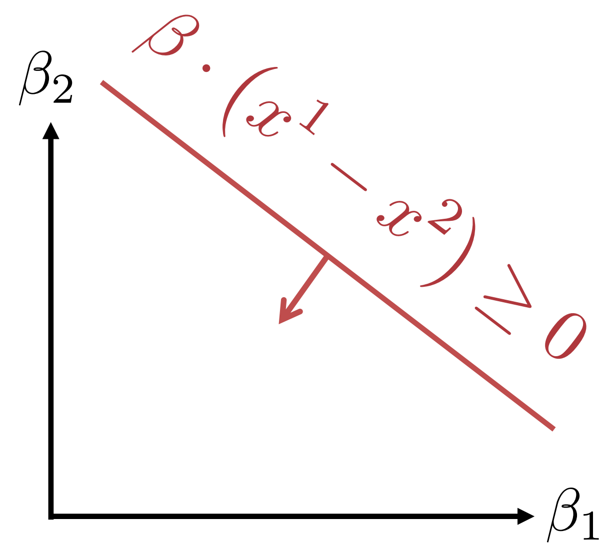
part-worths

product profile

noise (gumbel)

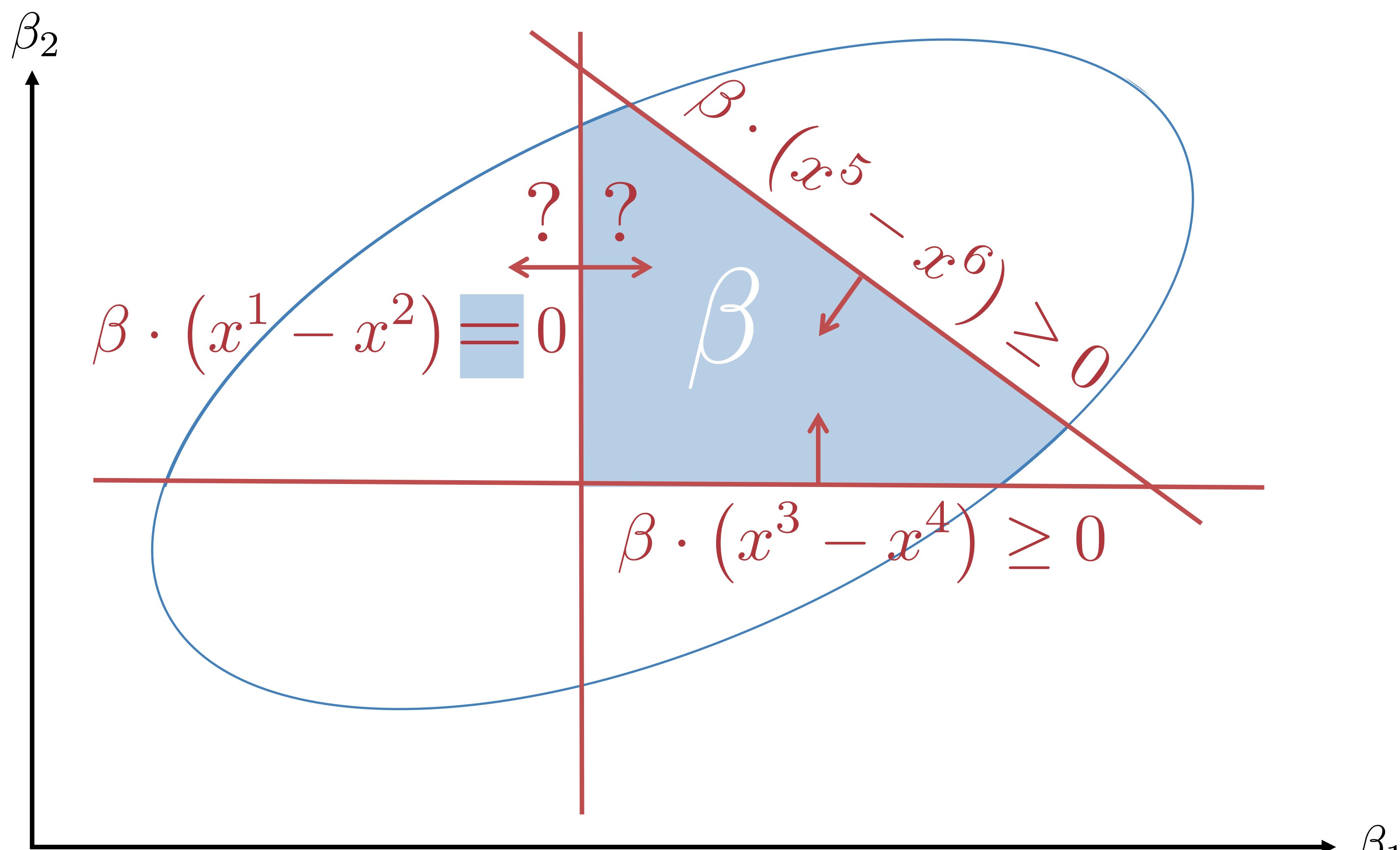
- Utility maximizing customer
  - Geometric interpretation of preference for product 1 without error

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



Polyhedral Method: Ask Question and Update

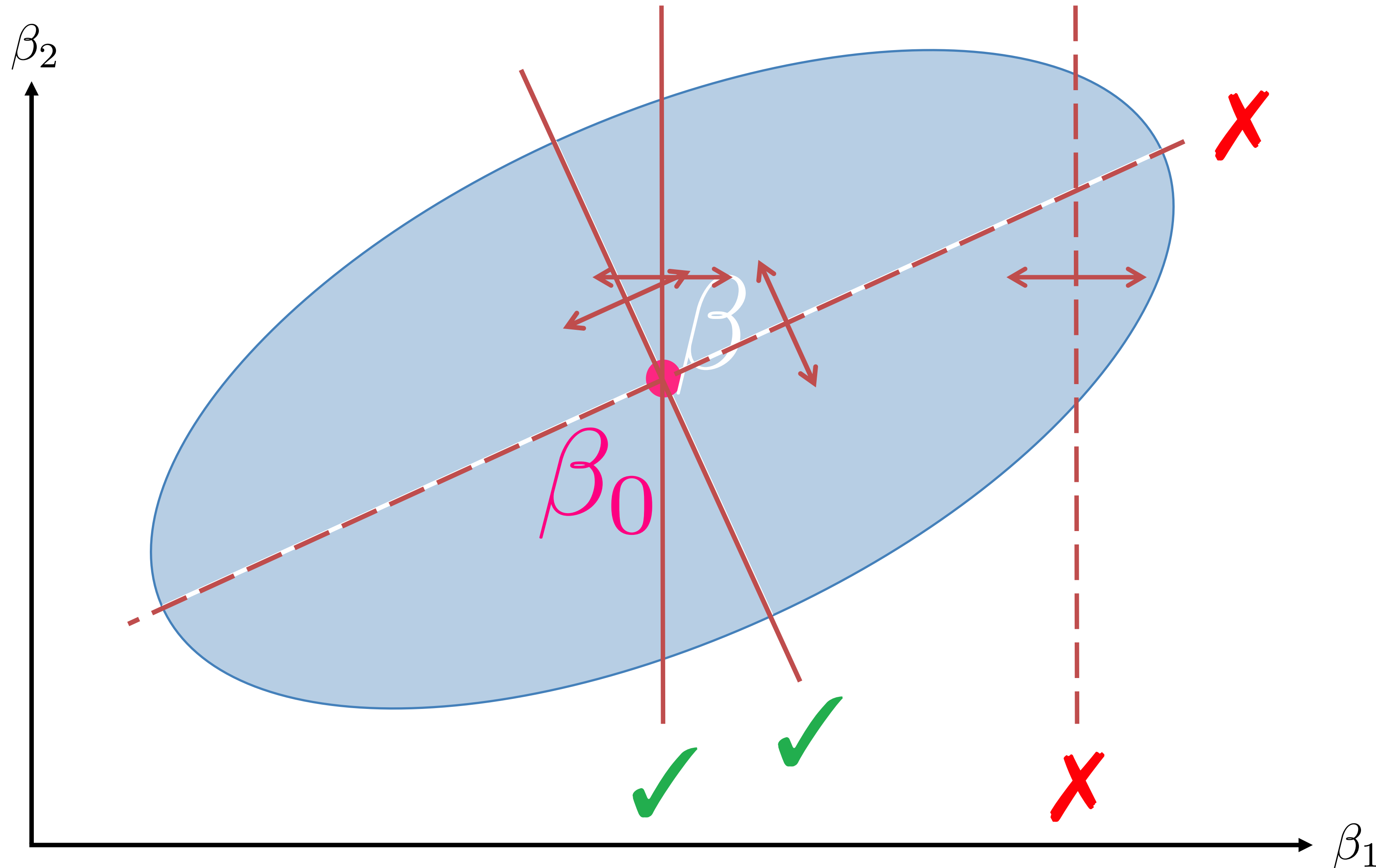
Geometric prior for β $\longrightarrow$ $x^1 \succeq x^2$ $\longrightarrow$ 3rd geometric posterior for β



Polyhedral: Estimation and Question Selection

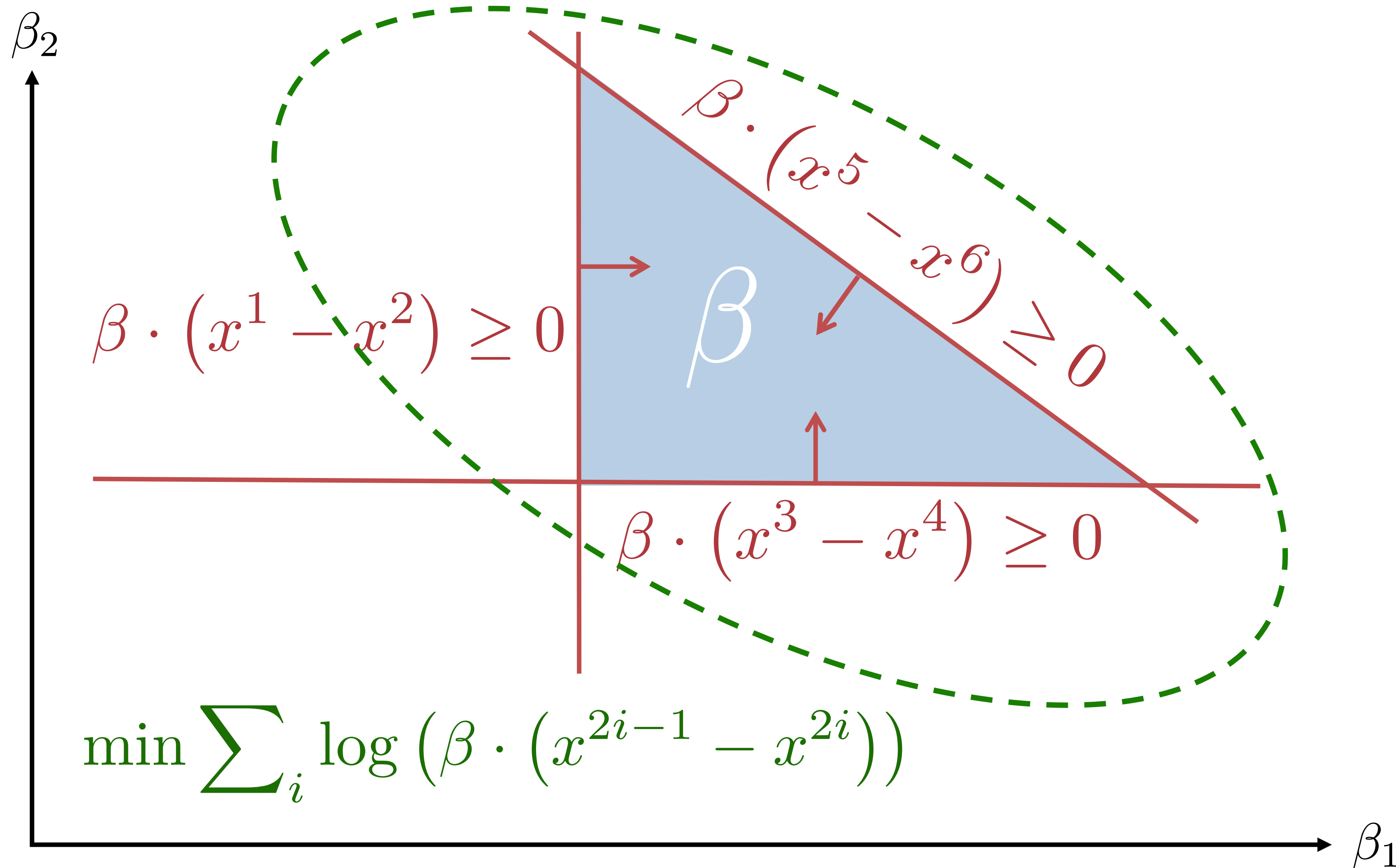
Good estimator for β? / Question?

1. Centroid ellipsoid / 2. Robust to bad geometry / Closest to β_0



Polyhedral Method: Non-ellipsoidal Sets

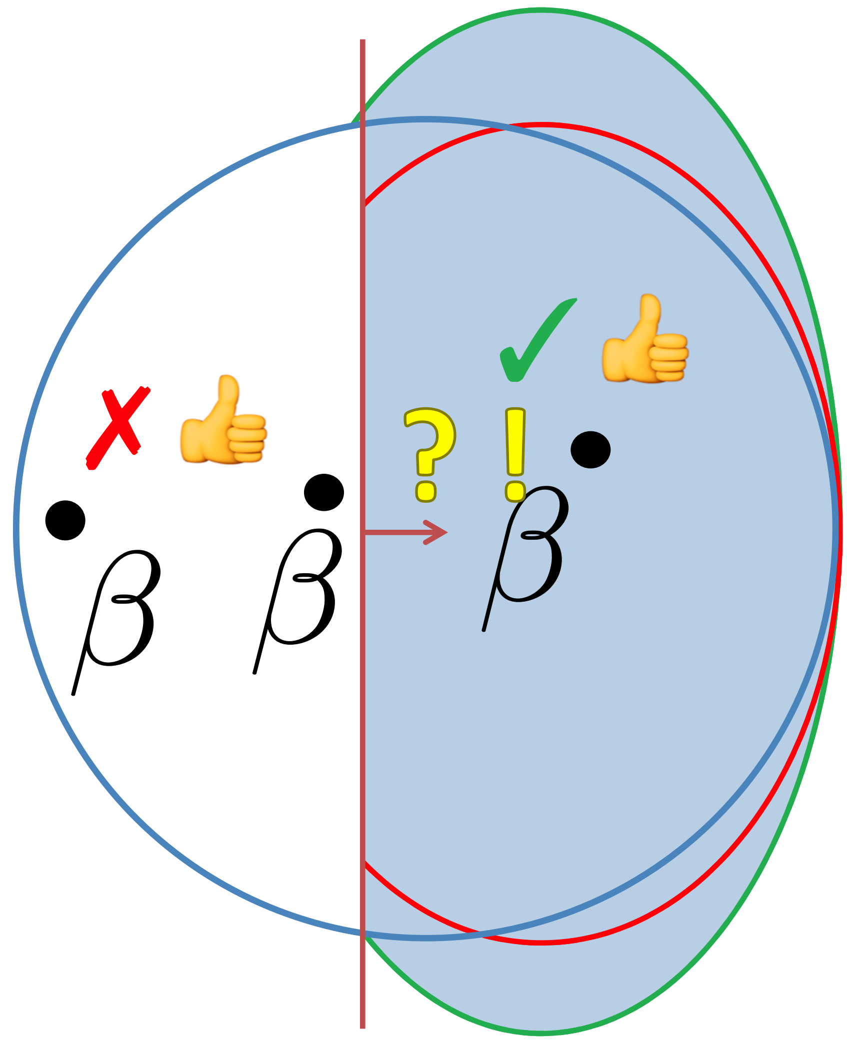
Idea from Nonlinear Programming (NLP): Approximate ellipsoid through analytic center.



Incorporating Response Error

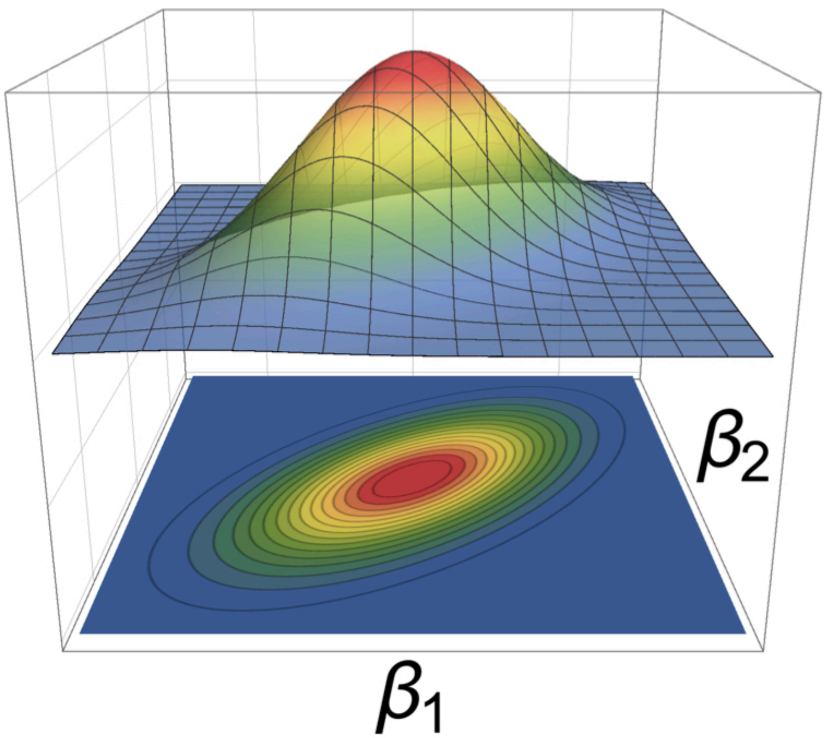
First Improvement: Ellipsoidal Updates

- 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



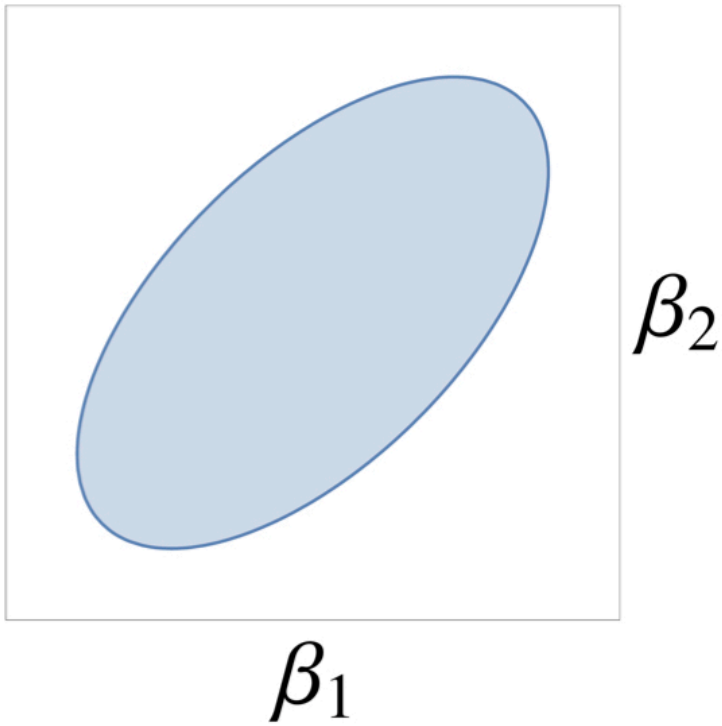
Distributions and Credibility Ellipsoids

Prior distribution of β



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

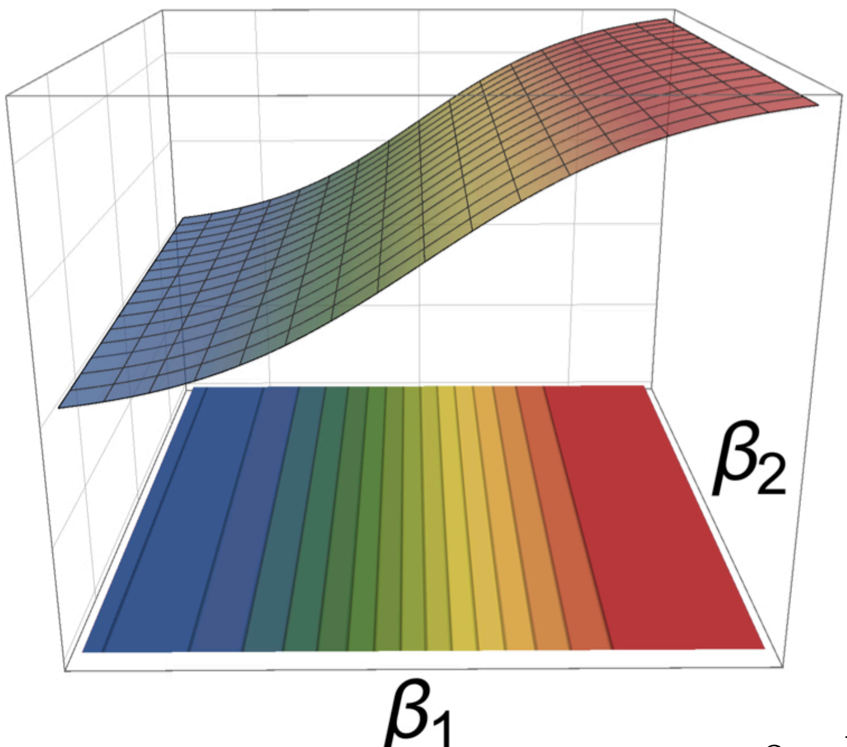
90% confidence/credibility ellipsoid



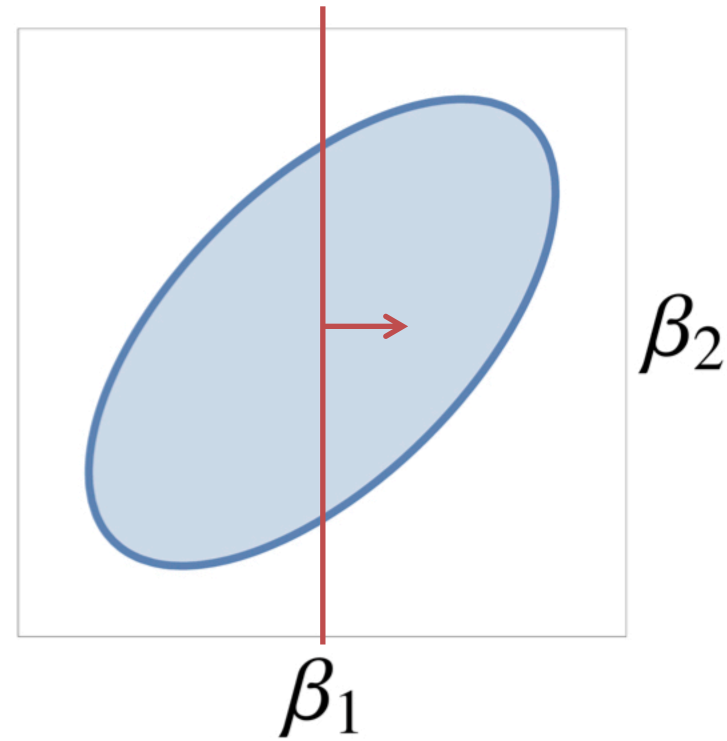
$$(\beta - \mu)' \cdot \Sigma^{-1} \cdot (\beta - \mu) \leq r$$

Answers with Error: Logit Probabilities

Likelihood Function



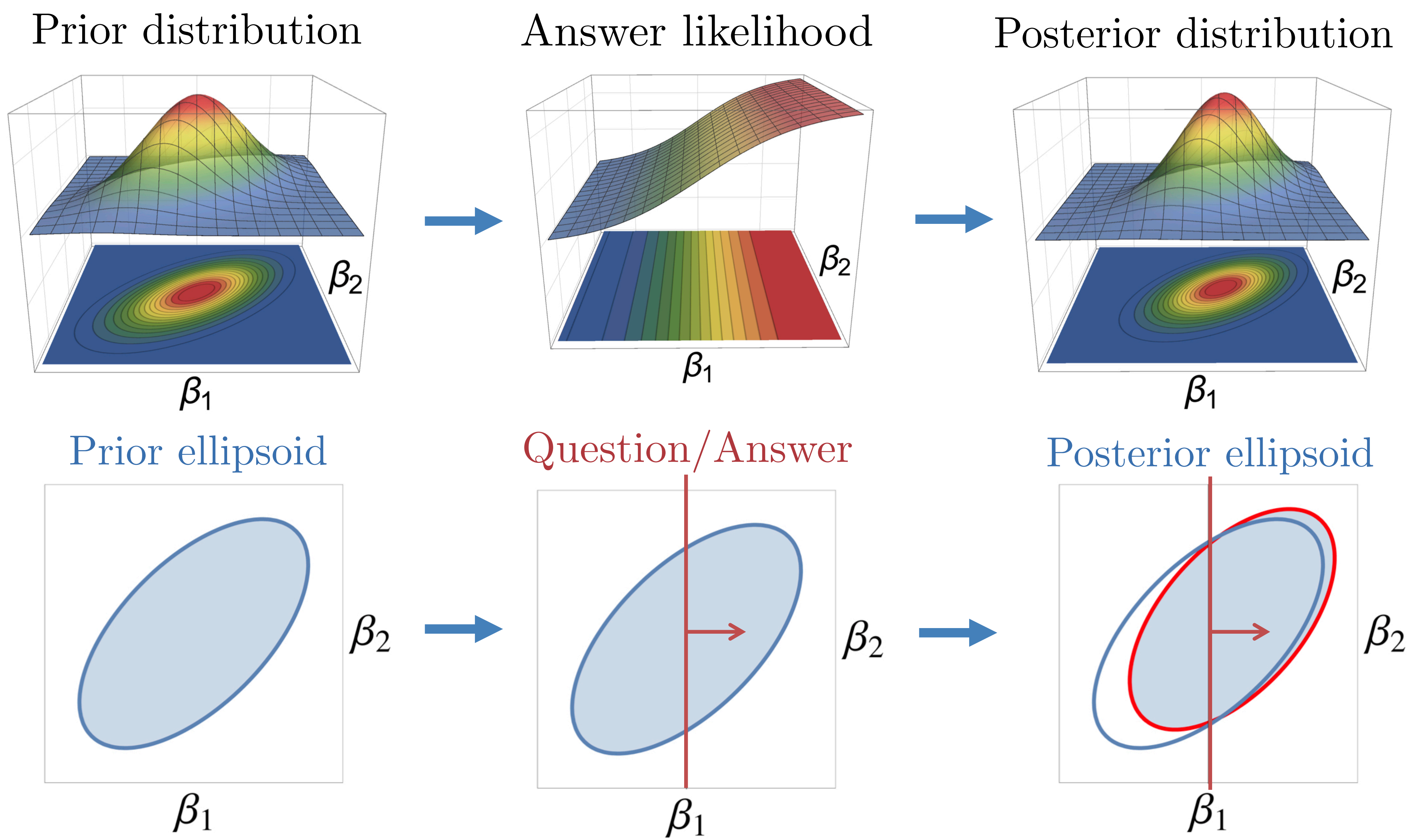
Question/Answer



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

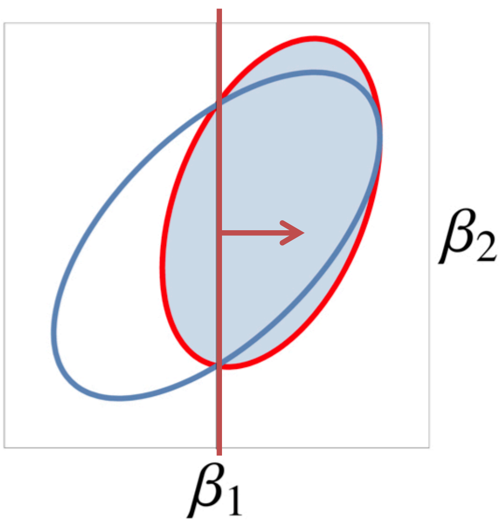
$$x^1 \succeq x^2$$

Bayesian Update and Geometric Updates



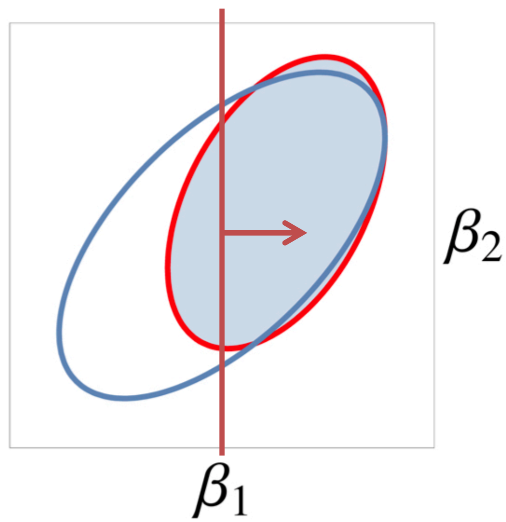
Geometric Comparison of Updates

Min. Volume Ellipsoid



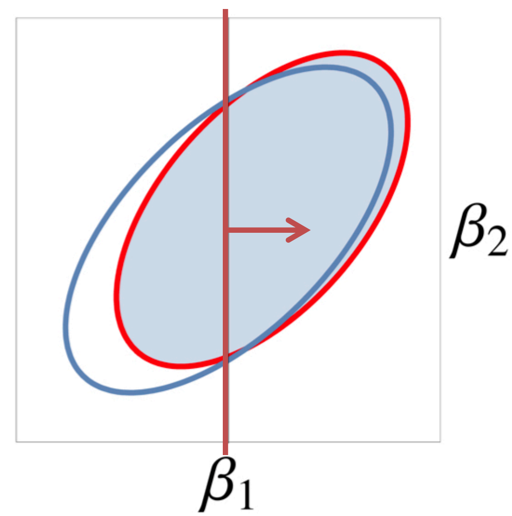
Simple Formula

Corrected Ellipsoid



Simple Formula

Bayesian for Normal Approx.



~~MCMC~~ 😐

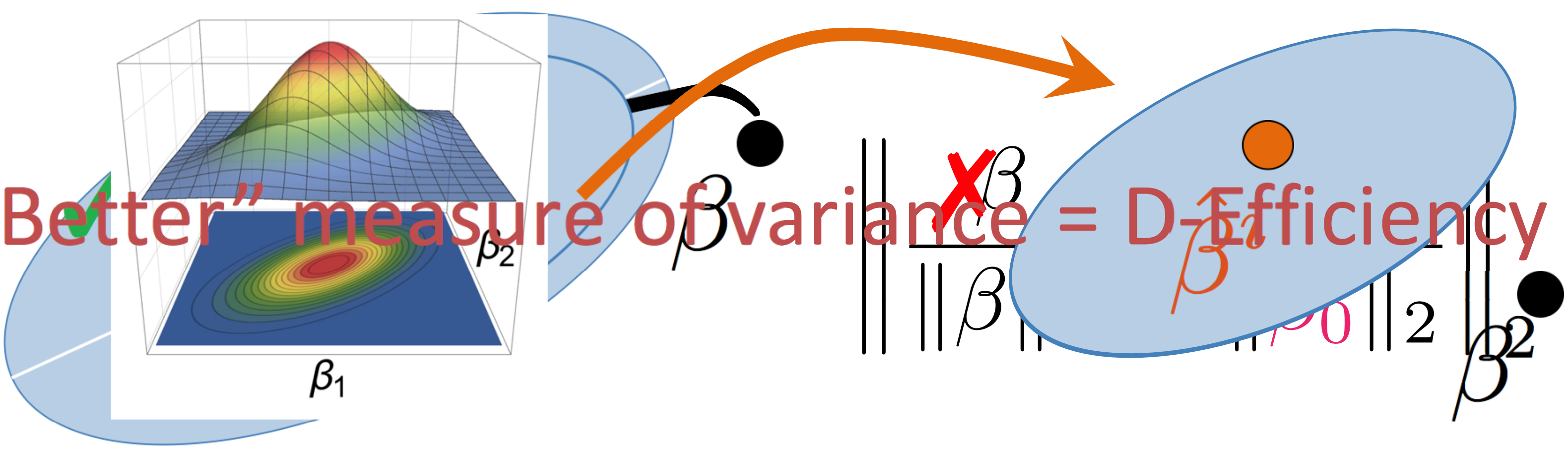
1-dim integral 🙂

$10^4 \to 10^7$ samples

Computational Comparison of Updates

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

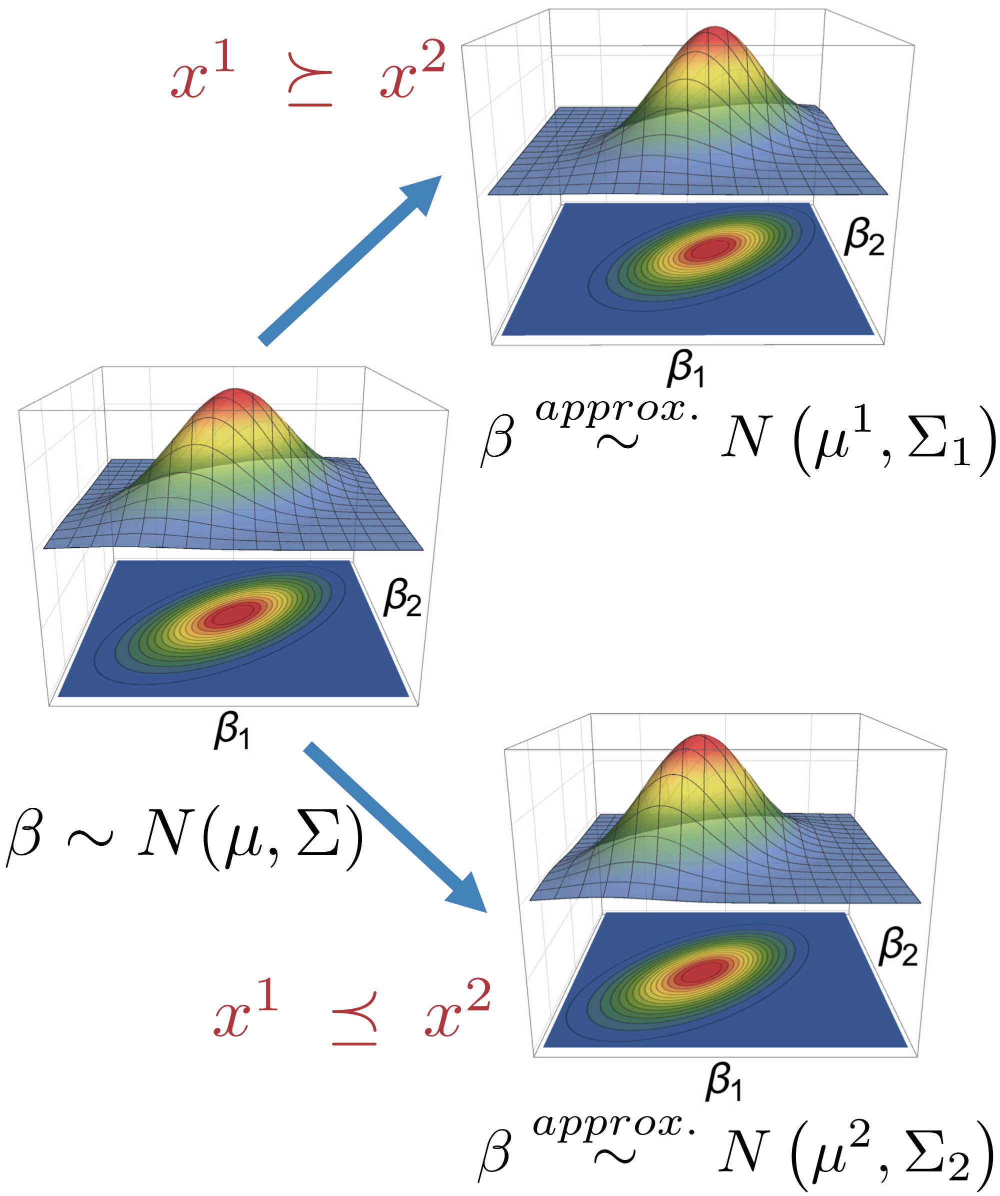
	Polyhedral	Ellipsoidal	Corrected Ellipsoidal	1-step Bayes
Feasible β	0.53	1	1	0.93
Distance (scaled)	0.92	0.86	0.88	0.85
Gaussian Volume	0.03	0.85	0.82	0.40



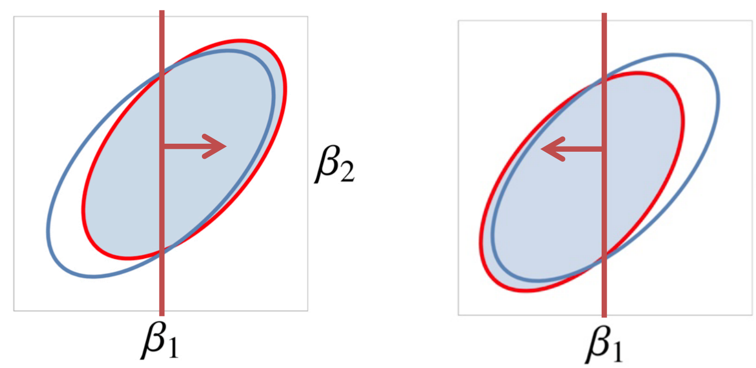
"Better" measure of variance = D-Efficiency

Improving Question Selection: Optimizing D-Efficiency

D-Efficiency and Posterior Covariance Matrix



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



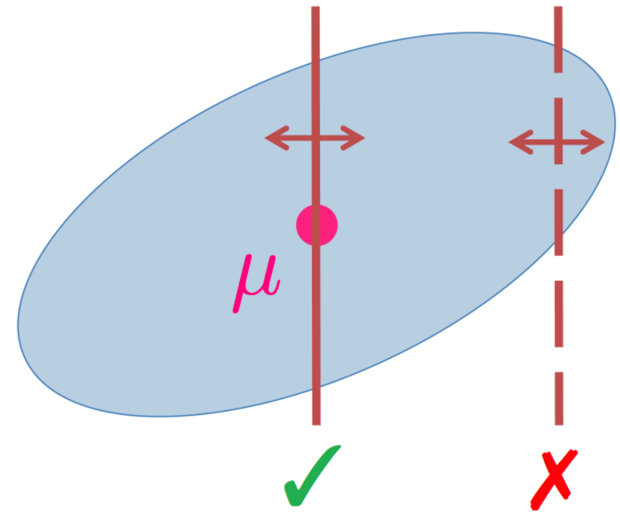
- Even evaluating expected D-Efficiency for a question requires multidimensional integration

Back to Question Selection: Property Trade-off

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

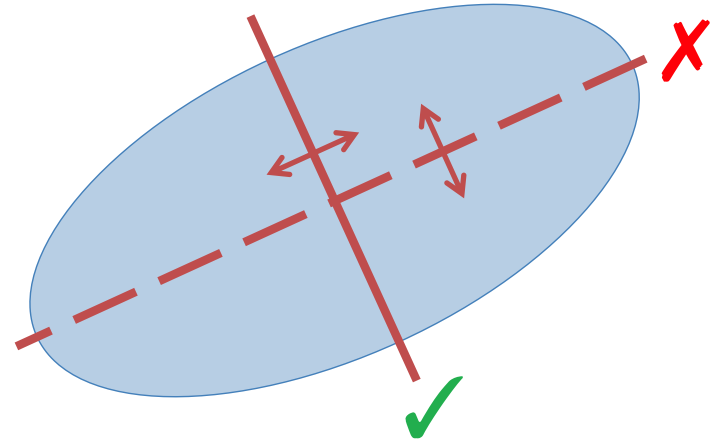
- Choice balance:
  - Minimize distance to center

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



- Postchoice symmetry:
  - Maximize variance of question

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

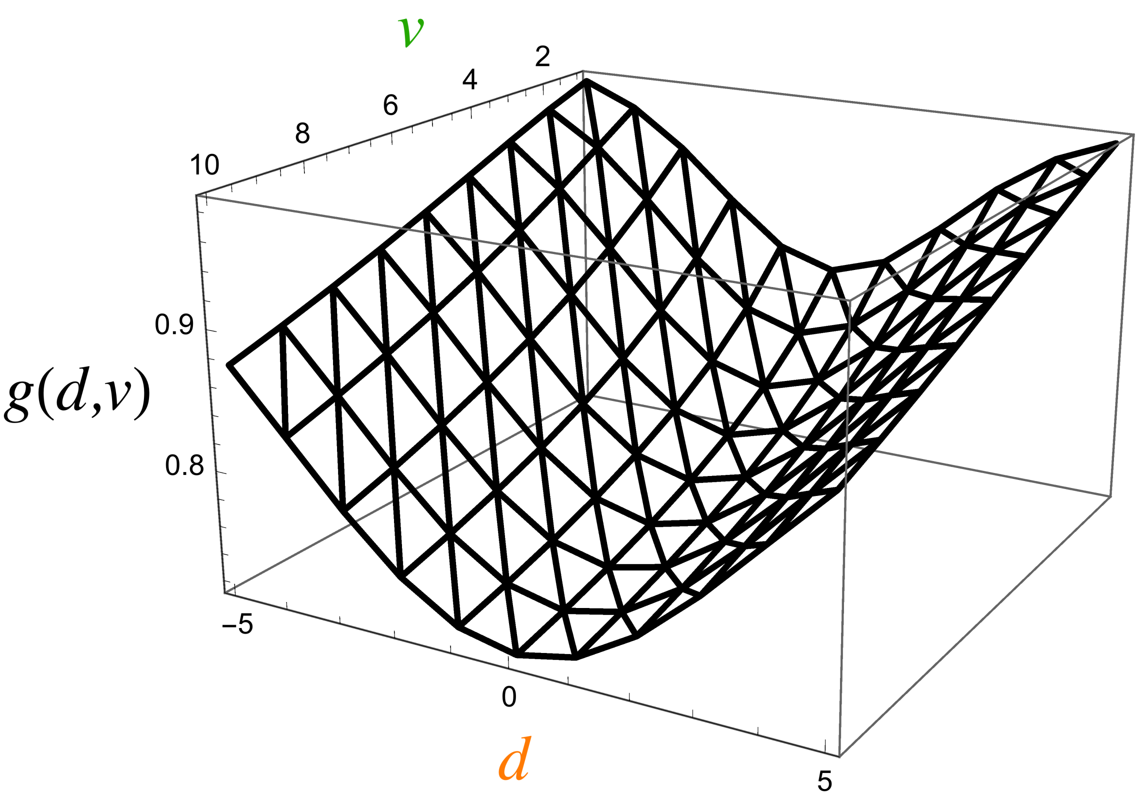


D-efficiency: Balance Question Trade-off

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

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

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



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

	Toubia et al.	PWL D-Efficiency
Feasible β	0.90	0.91
Distance (scaled)	0.97	0.85
D-Efficiency	2.2E+07	7.00E+06
Gaussian Volume	0.74	0.40

- 1 step for random covariance/ellipsoid

	Toubia et al.	PWL D-Efficiency
D-Efficiency	0.016	0.015
variance	110	83
distance	8.6	1.2
Area R1 / R2	32% / 68%	47% / 53%

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 and attribute levels
- Future:
  - Combination and comparison with fully Bayesian
  - Combine with polyhedral updates
  - Computational improvements
  - Field experiments