Ellipsoidal Methods for Adaptive Choice-based Conjoint Analysis

Juan Pablo Vielma

Massachusetts Institute of Technology

Workshop in Consumer Analytics,
San Pedro de Atacama, Chile. January, 2016.

Joint work Denis Sauré

Choice-based Conjoint Analysis (CBCA)

Feature	Chewbacca	BB-8
Wookiee	Yes	No
Droid	No	Yes
Blaster	Yes	No
I would buy toy	☑	☐

Product Recommendations via CBCA

- Very few questions (5) and possibly levels (2) and products per question (2):
  - Need very accurate question selection = adaptive
  - Need fast question selection ≠ full hierarchical bayes
- Good starting candidate = Polyhedral Method (Toubia et al. 2004)
  - Geometric/Bayesian interpretation
    - We improve update = geometric and quick bayes
    - We improve question selection = Mixed Integer Programming
    - We re-interpret question selection criteria = D-Efficiency

Choice-based Conjoint Analysis

Feature	Chewbacca	BB-8
Wookiee	Yes	No
Droid	No	Yes
Blaster	Yes	No
I would buy toy	☑	☐

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

Product Profile	x^1	x^2

Preference Model and Geometric Interpretation

- Utilities for 2 products, d features, logit model

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

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

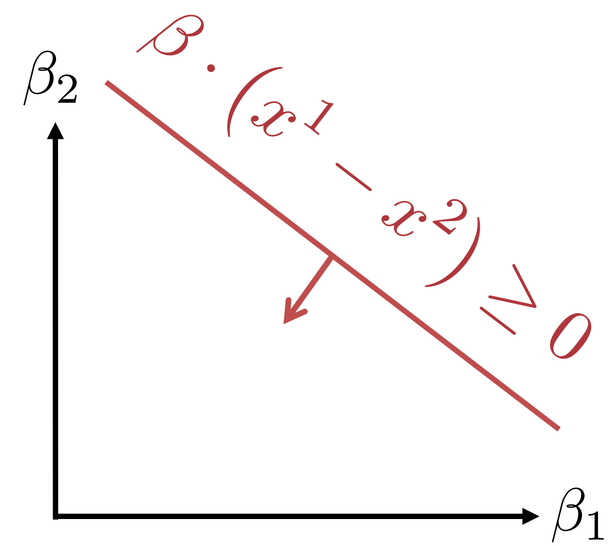
part-worths

product profile

noise (gumbel)

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

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

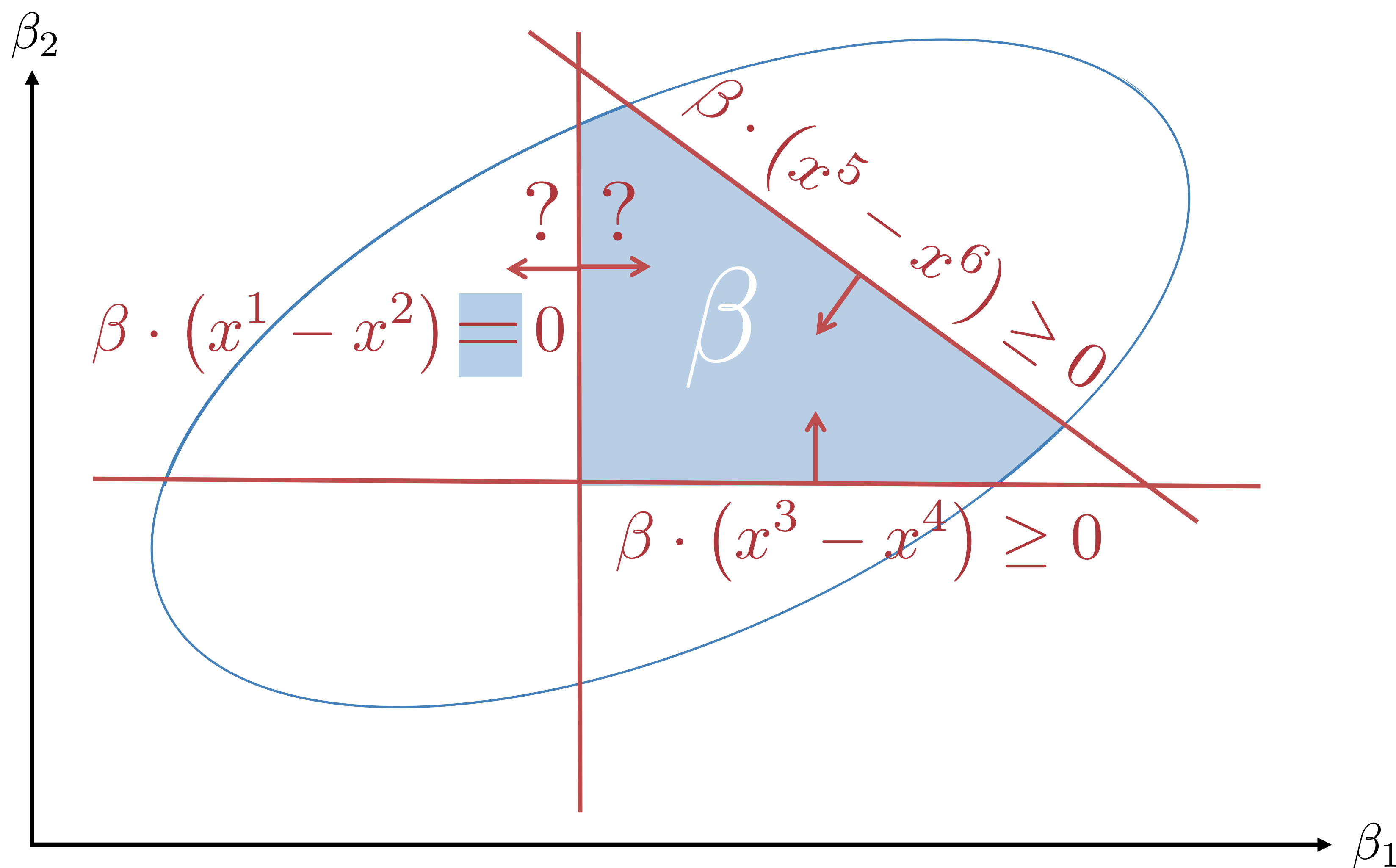


Polyhedral Method (Toubia et al. 2004)

- Pros:
  - Very elegant purely geometric method
  - No sampling required = very quick
  - Output is convex set that contains part-worth vector β
    - Point estimation from set or risk-aware robust optimization
  - Very good for high heterogeneity and low question error
- Cons:
  - Does not consider question error
    - Patches by Toubia et al. 2007 and Bertsimas O'Hair 2013, but loose elegance, interpretability and simplicity
  - Question selection is good, but heuristic (can fail)

Polyhedral Method: Ask Question and Update

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



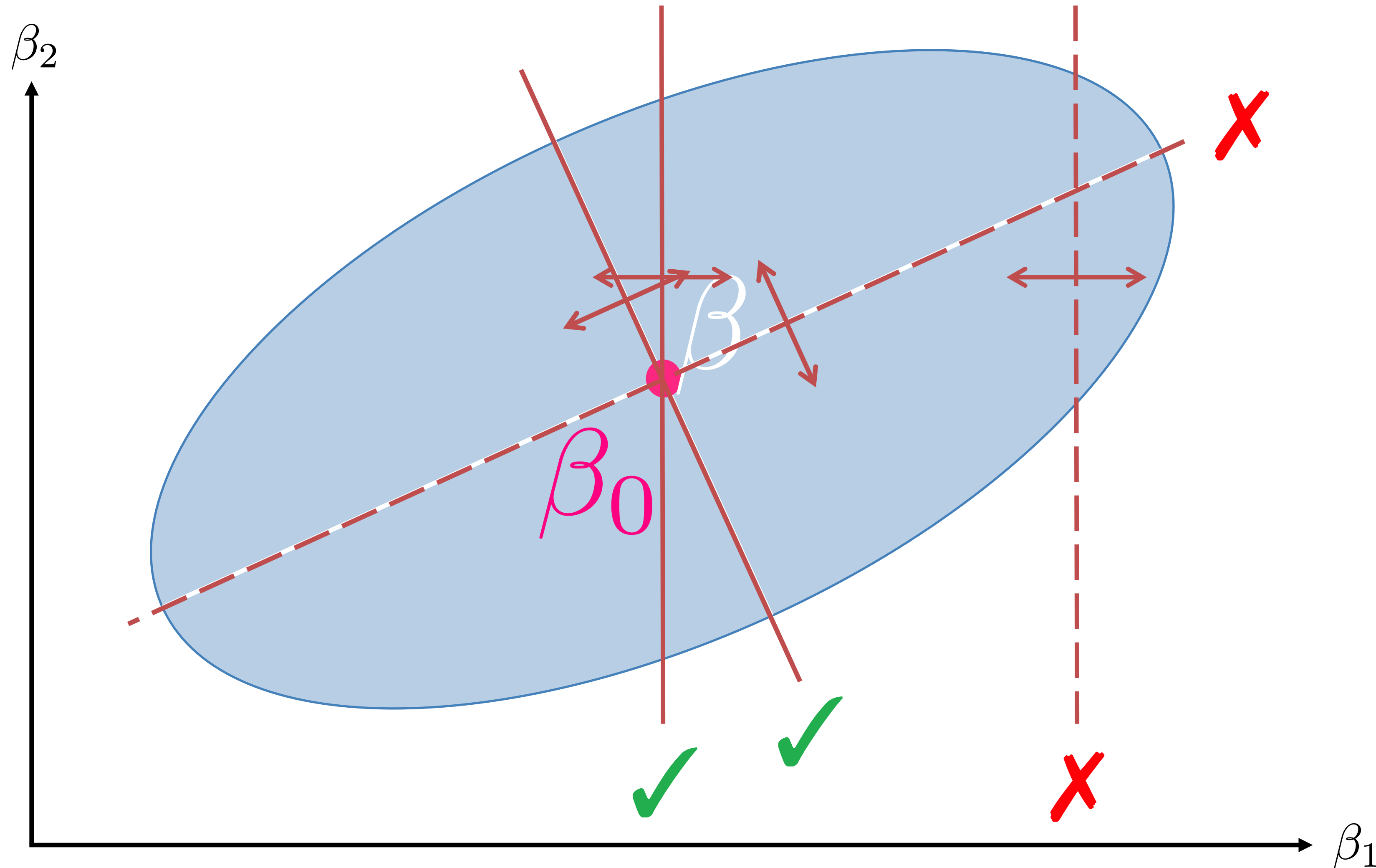
Polyhedral: Estimation and Question Selection

Good Question?

Good estimator for β?

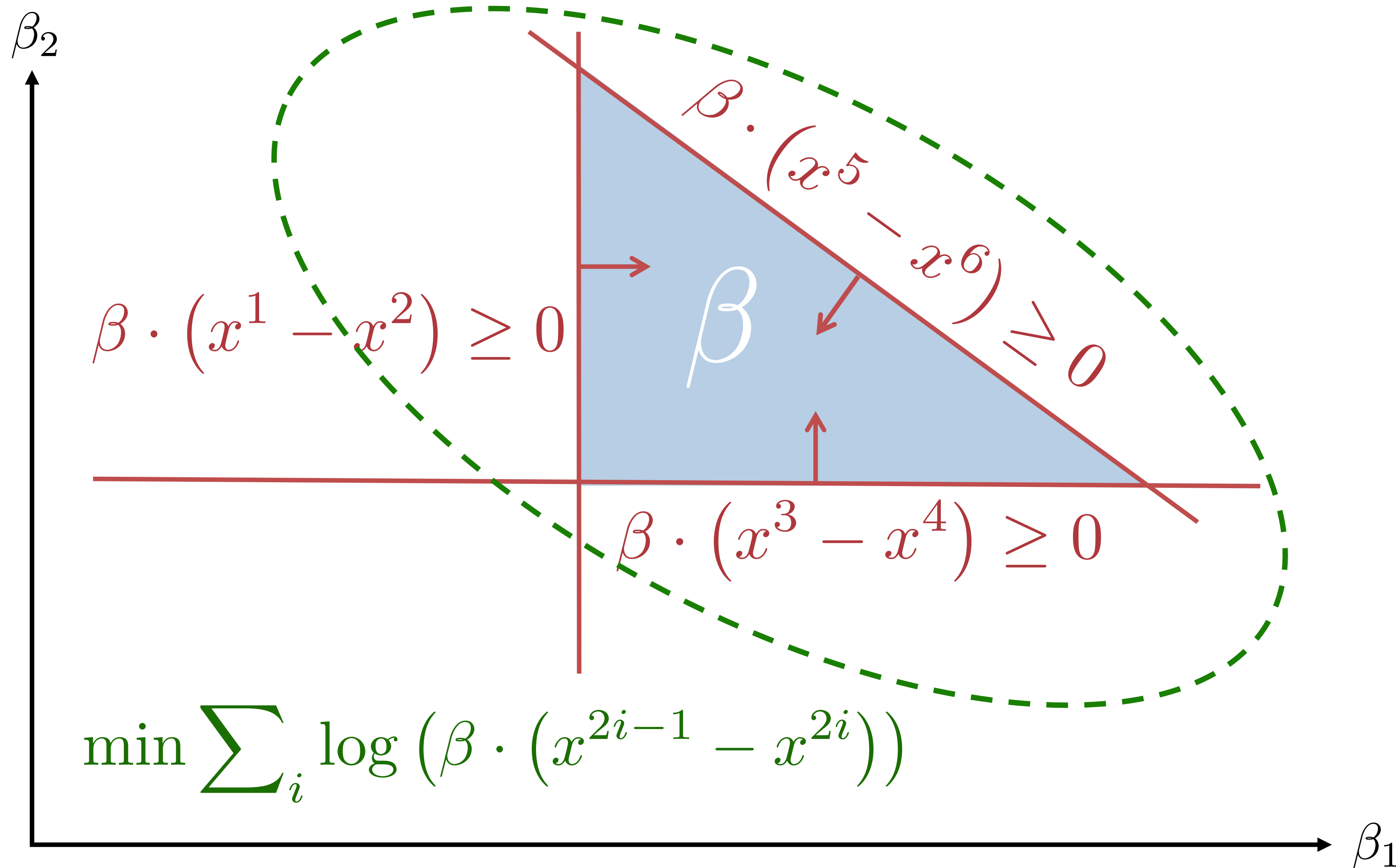
1. Center of ellipsoid

2. Robust: symmetry β_0



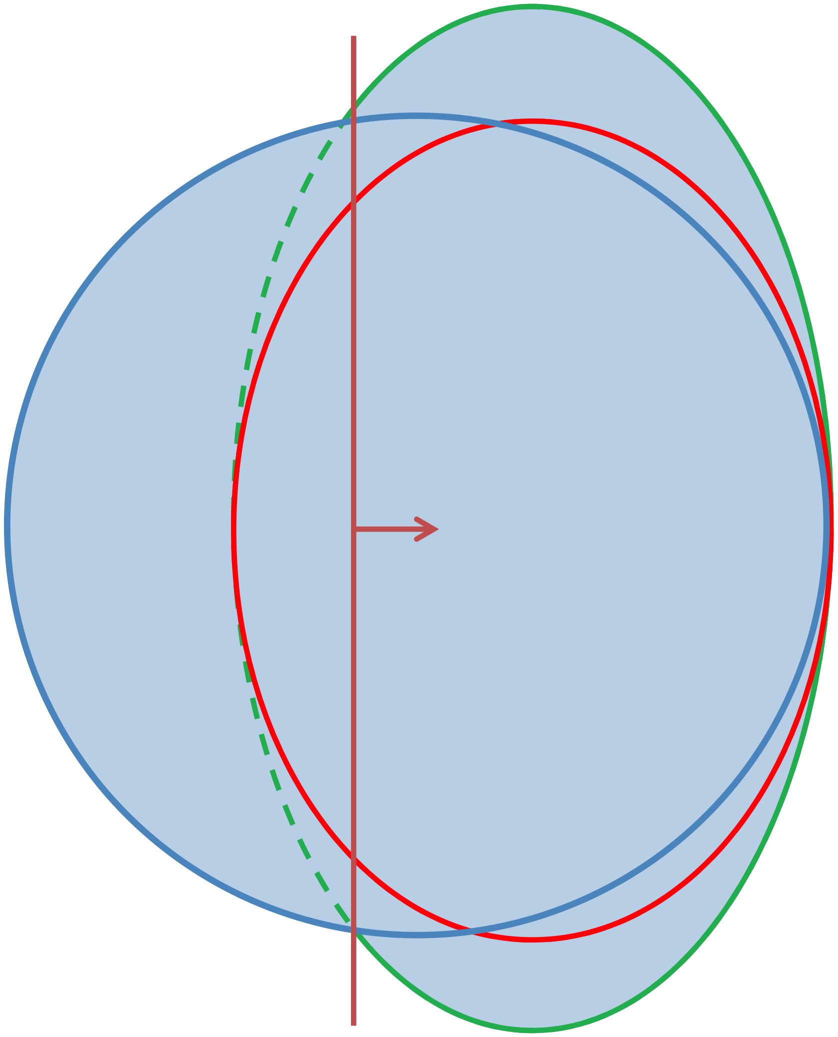
Polyhedral Method: Non-ellipsoidal Sets

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



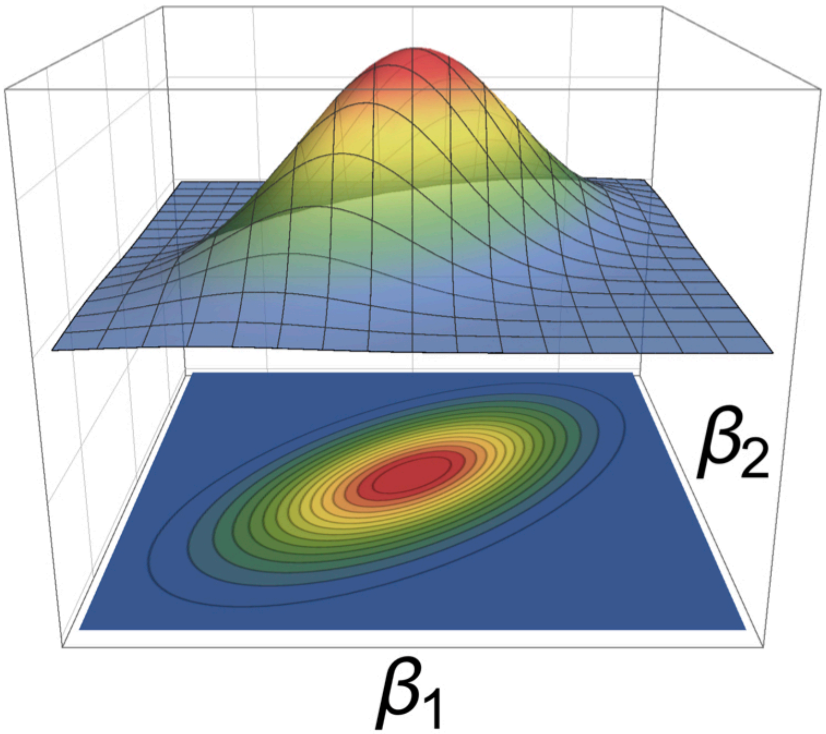
First Improvement: Ellipsoidal Updates

- 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



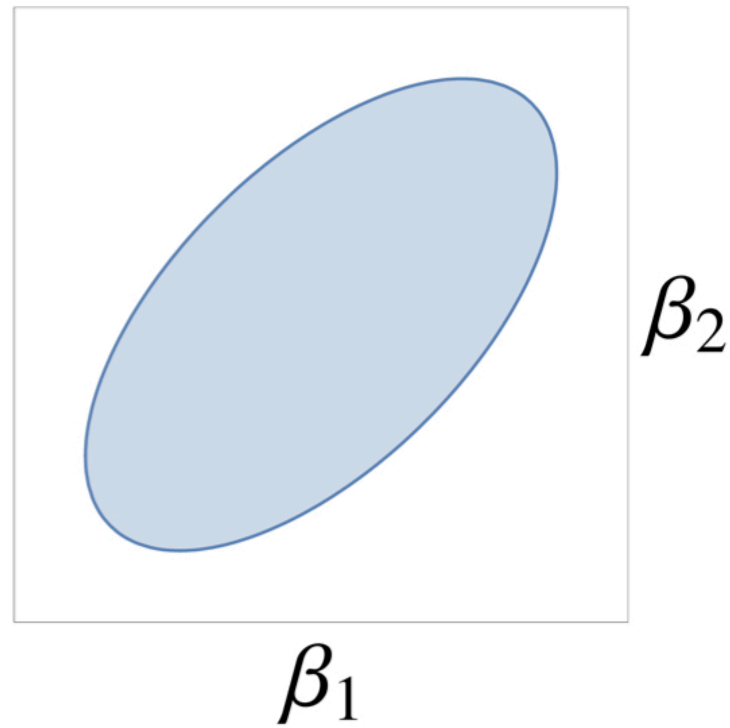
Distributions and Credibility Ellipsoids

Prior distribution of β



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

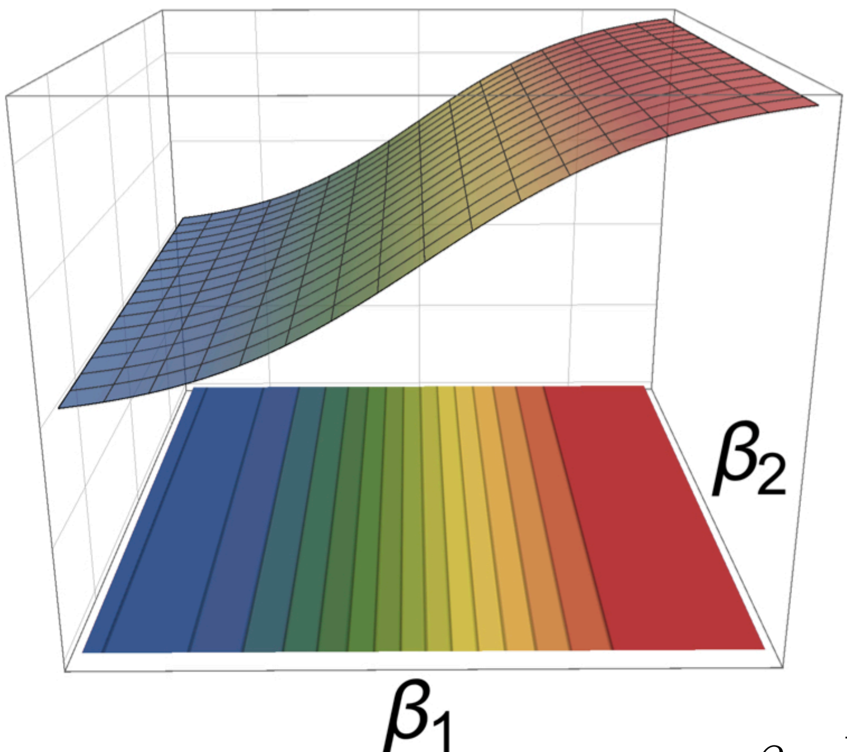
90% confidence/credibility ellipsoid



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

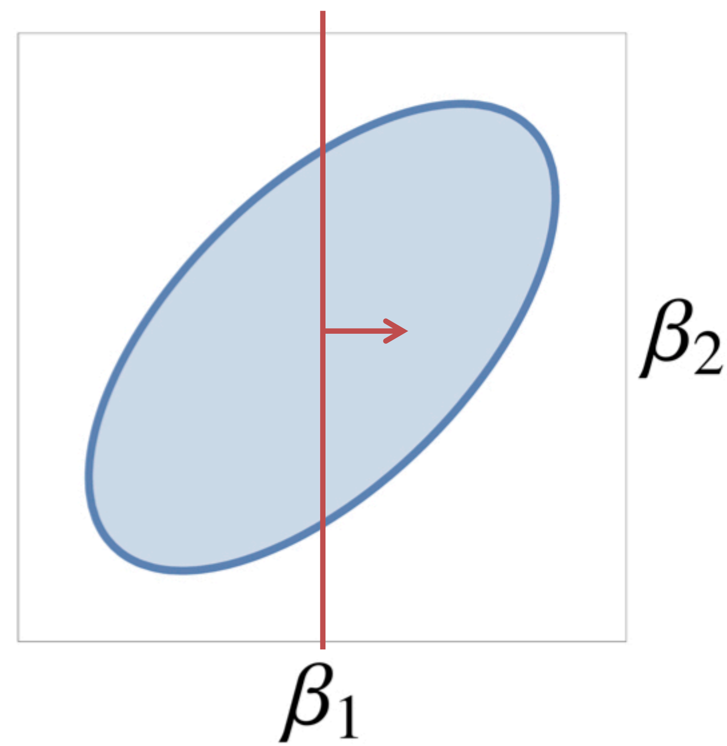
Answers with Error: Logit Probabilities

Likelihood Function



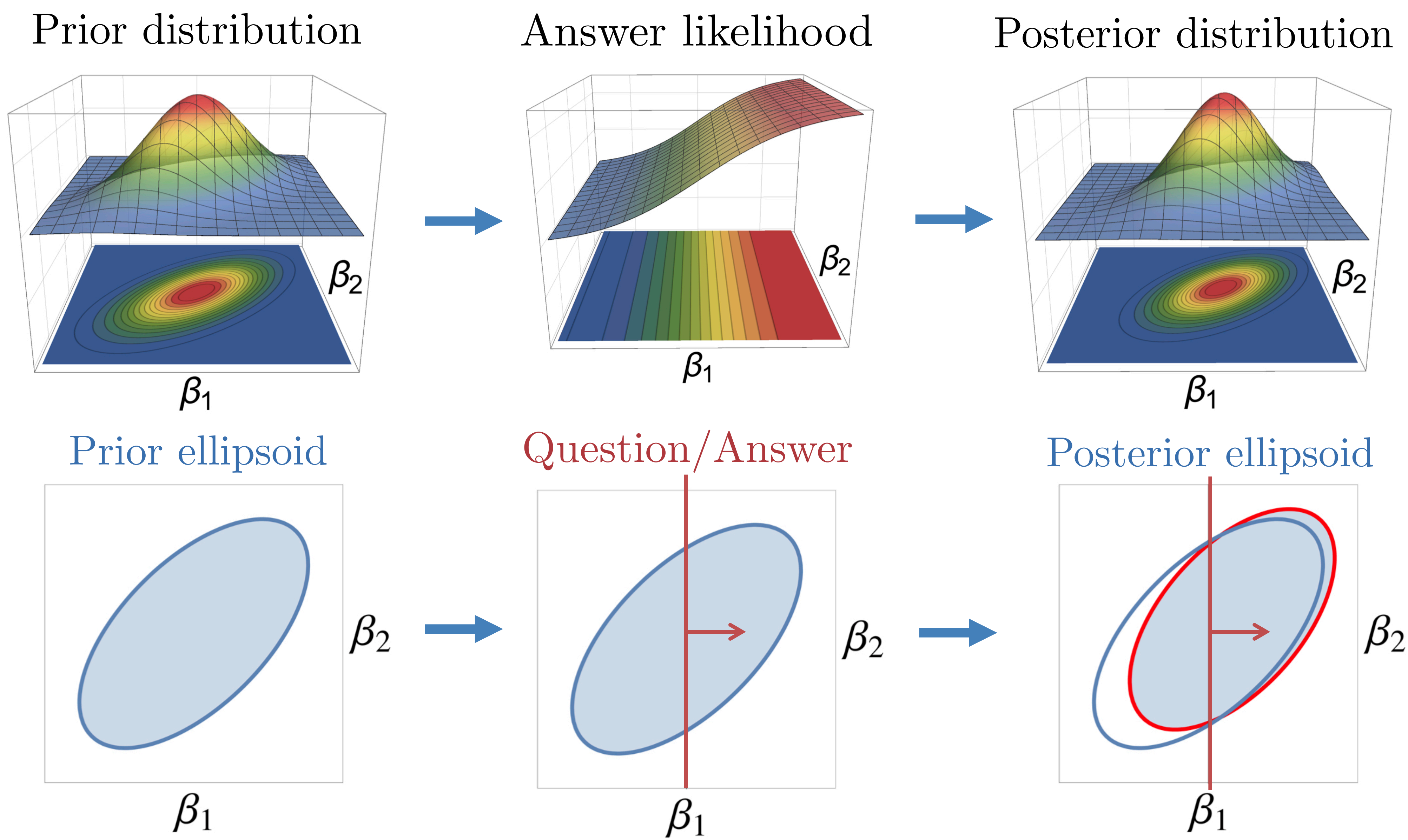
$$\mathbb{P}\left(x^1 \succeq x^2 \mid \beta\right) = \frac{e^{\beta \cdot x^1}}{e^{\beta \cdot x^1} + e^{\beta \cdot x^2}}$$

Question/Answer



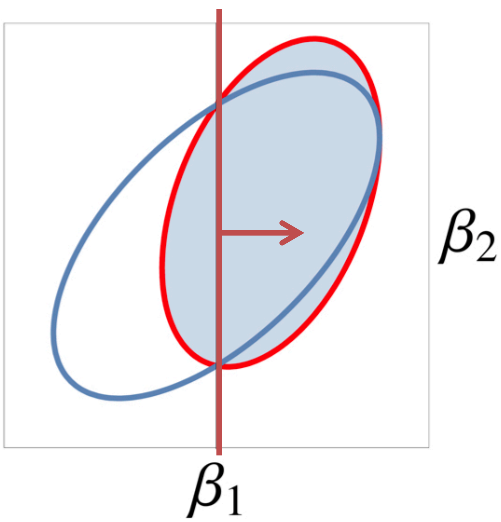
$$x^1 \succeq x^2$$

Bayesian Update and Geometric Updates



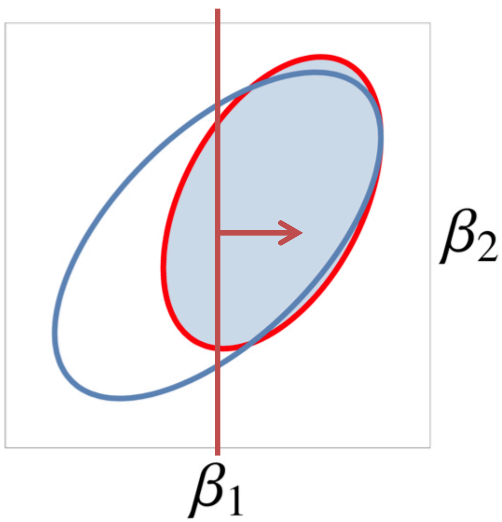
Geometric Comparison of Updates

Min. Volume Ellipsoid



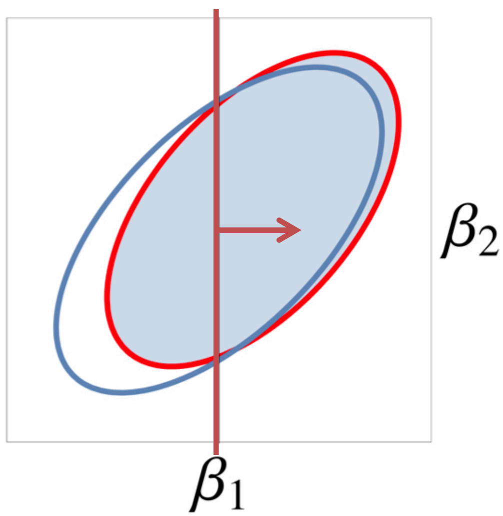
Simple Formula

Corrected Ellipsoid



Simple Formula

Bayesian for Normal Approx.

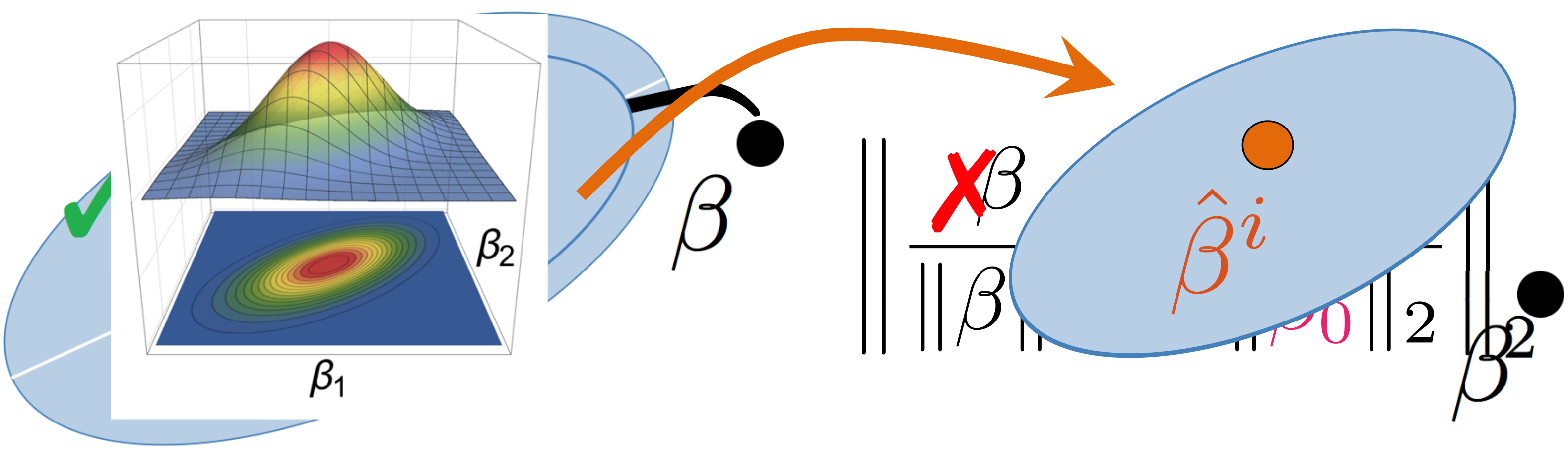


1-dim integral

Computational Comparison of Updates

- Gaussian prior and 90% credibility ellipsoid
  - 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

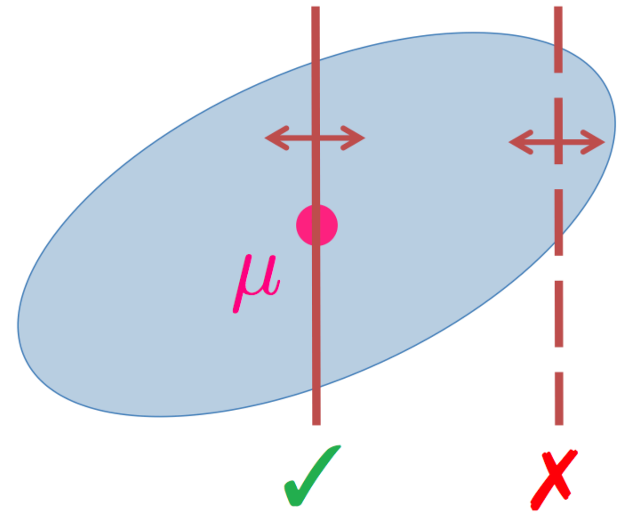


Back to Question Selection: Property Trade-off

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

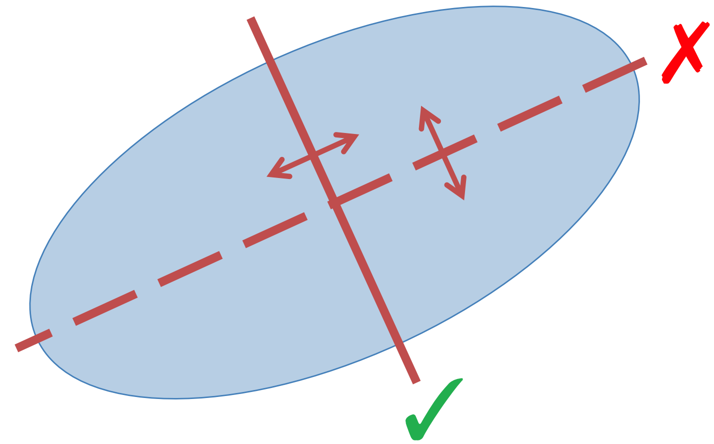
- Choice balance:
  - Minimize distance to center

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



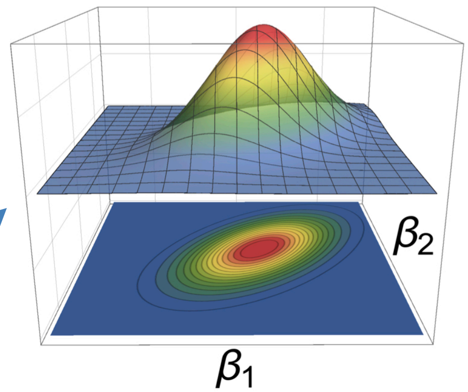
- Postchoice symmetry:
  - Maximize variance of question

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

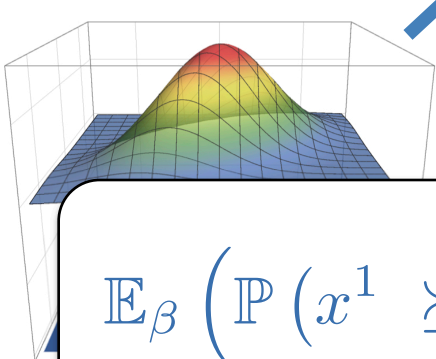


D-Efficiency and Posterior Covariance Matrix

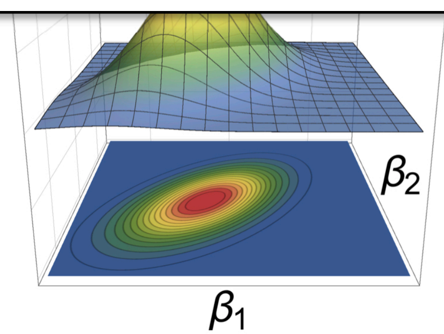
$$x^1 \succeq x^2$$



$$\beta \overset{approx.}{\sim} N\left(\mu^1, \Sigma_1\right)$$



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

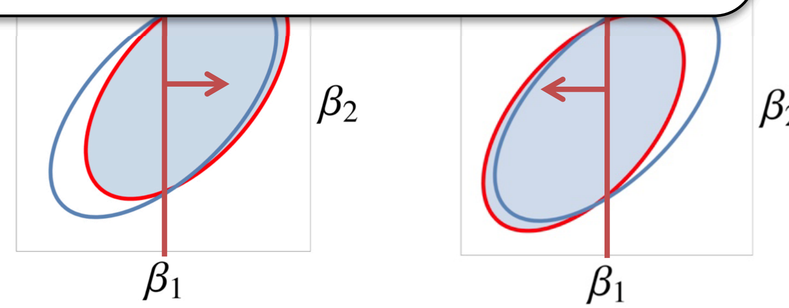


$$x^1 \preceq x^2$$

$$\beta \overset{approx.}{\sim} N\left(\mu^2, \Sigma_2\right)$$

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

$$\mathbb{E}_\beta\left(\mathbb{P}\left(x^1 \succeq x^2 \mid \beta\right)\det(\Sigma_1)^{1/p} + \mathbb{P}\left(x^1 \preceq x^2 \mid \beta\right)\det(\Sigma_2)^{1/p}\right)$$

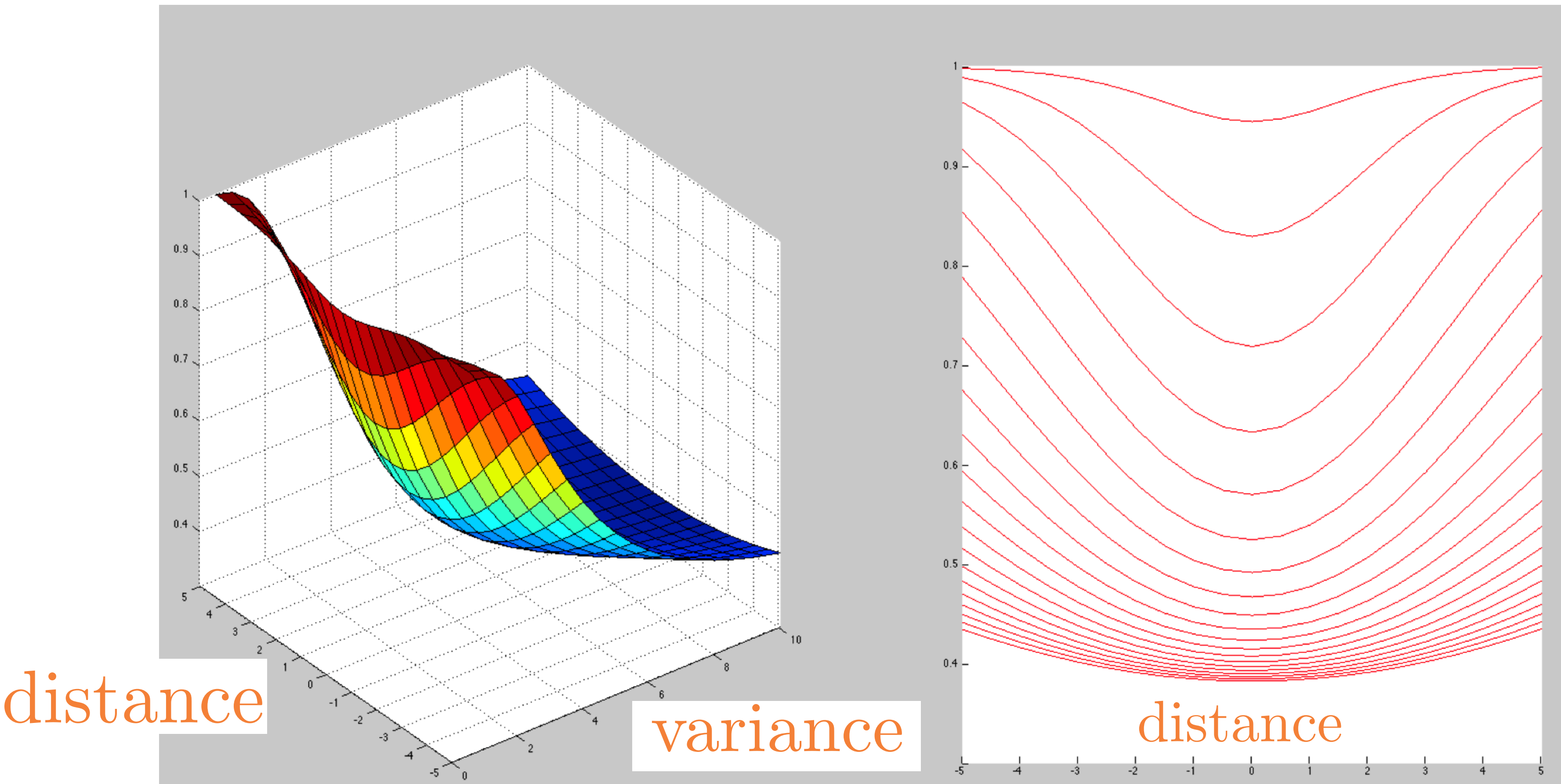


D-efficiency: Balance Question Trade-off

- D-efficiency = Nonlinear function of

$$\text{distance} = \mu \cdot (x^1 - x^2)$$

$$\text{variance} = (x^1 - x^2)' \cdot \Sigma \cdot (x^1 - x^2)$$



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	0.88	0.92
Distance (scaled)	0.99	0.86
D-Efficiency	2.3E+07	7.01E+06
Gaussian Volume	0.75	0.40

- 1 step for random covariance/ellipsoid

	Toubia et al.	PWL D-Efficiency
Balance	0.39	0.83
D-Efficiency	0.018	0.016
distance	9.44	0.17
variance	111	80

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
- Future:
  - Combination and comparison with fully Bayesian
  - Pre-computing and Real-Time
  - Use for recommendation