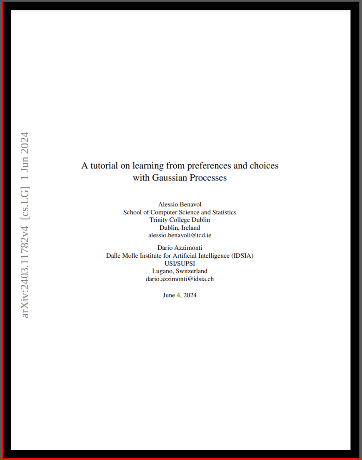
Learning from Preferences and Choice functions

Alessio Benavoli
Associate Professor
School of Computer Science and Statistics
Trinity College Dublin
alessio.benavoli@tcd.ie

SIPTA Seminar 24 October 2024

Background



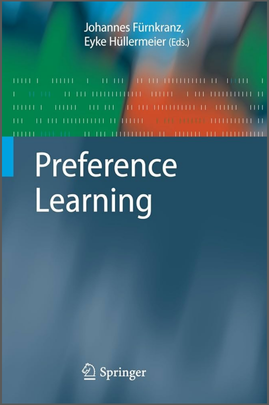
arXiv:2403.11782v4 [cs.LG] 1 Jun 2024

A tutorial on learning from preferences and choices
with Gaussian Processes

Alessio Benavol
School of Computer Science and Statistics
Trinity College Dublin
Dublin, Ireland
alessio.benavoli@tcd.ie

Dario Azzimonti
Dalle Molle Institute for Artificial Intelligence (IDSIA)
USI/SUPSI
Lugano, Switzerland
dario.azzimonti@idsia.ch

June 4, 2024



Outline

Learning from Human Feedback

Object Preference

Consistent Preferences
Less rationality
- discernibility (Luce)
- noise (Thurstone)
More rationality
- Monotonicity

Label Preference

- Plackett-Luce Model
- Thurstone's Model
- Bradley-Terry Model

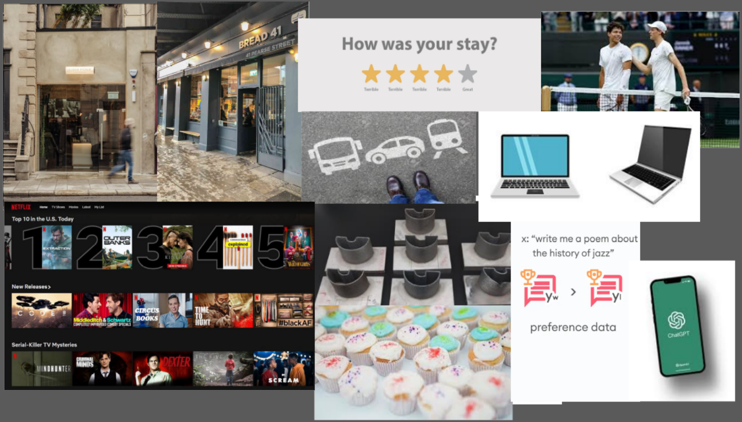
Choice functions

- Pareto rationality
- quasi-rationality

Utility function

Two-argument function

Preferences



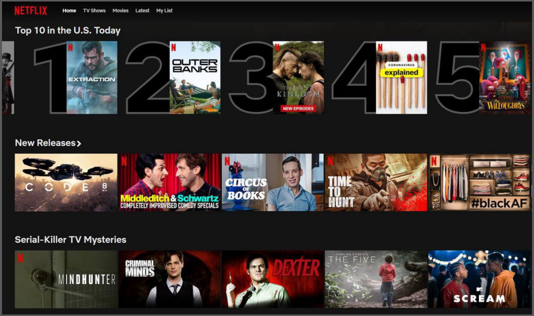
PL aims to learn preference models from observed, revealed or automatically extracted preference information.

Object Preference

Laptop Model	Laptop 1	Laptop 2
Processor	Intel I7	Intel I7
Generation	10th	11th
Operating System	Windows	MacOS

- a set of objects $\{k |\ k = 1, \ldots, m\}$
- a vector of characteristics for each objects $\mathbf{x}_k \in \mathbb{R}^q$
  - $\mathcal{X} = \{\mathbf{x}_k :\ \ k = 1, \ldots, m\}$
- an associated set of pairwise preferences of the form $\mathbf{x}_i \succ \mathbf{x}_j$

Label Preference



- a set of instances $\mathcal{X} = \{\mathbf{x}_k : \; k = 1, \ldots, m\}$
- a set of labels $\mathcal{Y} = \{y_1, \ldots, y_\ell\}$
- for each instance $\mathbf{x}_k$ an associated set of pairwise preferences of the form $y_i \succ_{\mathbf{x}_k} y_j$

What is a (strict) preference?

Asymmetry

If  $\succ$  then  $\nsucc$ 

Negative Transitivity

If  $\succ$  then for any 

either  $\succ$  or  $\succ$  or both

A strict preference relation is said to be *consistent* – *rational* – when it satisfies the above two properties.

Utility representation

Definition

For a preference relation $\succ$ on $\mathcal{X}$, the function $u : \mathcal{X} \to \mathbb{R}$ represents $\succ$ if

$$\mathbf{x}_i \succ \mathbf{x}_j \quad \text{if} \quad u(\mathbf{x}_i) > u(\mathbf{x}_j). \tag{1}$$

We say that u is a utility function for $\succ$.

Example

utility=taste

utility=performance

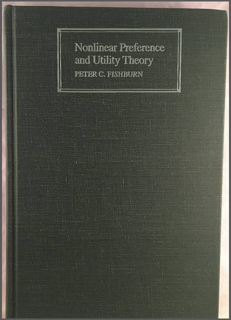
Theorem (Debreu 1954)

The relation $\succ$ admits a utility function representation iff it is **consistent** *(asymmetric and negatively transitive)*

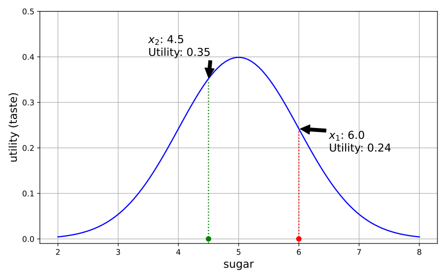
Two-argument function representation

Definition

$$\mathbf{x}_i \succ \mathbf{x}_j \quad \textit{if} \quad q(\mathbf{x}_i, \mathbf{x}_j) > 0 \tag{2}$$



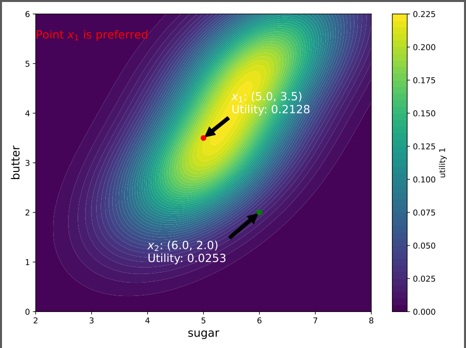
Example



4.5  $\succ$ 6 

$\Big\{$ 4.5  $\succ$ 6 , 4  $\succ$ 2.5 , ... $\Big\} \Rightarrow u(sugar)$

Example



(5,3.5)  $\succ$ (6,2) 

$\Bigg\{$ (5,3.5)  $\succ$ (6,2)  $, \ldots \Bigg\} \Rightarrow u(sugar, butter)$

Object preference learning

1. Given $\mathcal{D}_m = \{x_1 \succ x_3, x_4 \succ x_2, x_5 \succ x_1, \ldots\}$
2. Predict $x_2 \overset{?}{\succ} x_3$

Object PL: Parametric

From Debreu's theorem, we can derive that

$$\mathcal{D}_m = \{\mathbf{x}_l^{(s)} \succ \mathbf{x}_r^{(s)} : \;\; s = 1, \ldots, m\} = \{u(\boldsymbol{x}_l^{(s)}) - u(\boldsymbol{x}_r^{(k)}) > 0 : \;\; s = 1, \ldots, m\}$$

Functional form for $u(\mathbf{x})$:

Linear [Har-Peled et al., 2002]: $u(\mathbf{x}) = \mathbf{w}^\top \mathbf{x}$

Log-Linear [Dekel et al., 2003]: $\ln u(\mathbf{x}) = \mathbf{w}^\top \mathbf{x}$

Non-linear (NNs): $u(\mathbf{x}) = \mathbf{w}^\top \mathbf{b}_\theta(\mathbf{x})$

The unknowns are determined by either solving a **constrained optimization** problem or by minimizing a loss function. No uncertainty representation.

Object PL: Non-parametric (Gaussian Process)

$$\mathcal{D}_m = \{\mathbf{x}_l^{(s)} \succ \mathbf{x}_r^{(s)} : \;\; s = 1, \ldots, m\} = \{u(\boldsymbol{x}_l^{(s)}) - u(\boldsymbol{x}_r^{(k)}) > 0 : \;\; s = 1, \ldots, m\}$$

By defining the objects as a matrix $X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r]^\top$ and the vector $\mathbf{u}(X) = [u(\boldsymbol{x}_1), u(\boldsymbol{x}_2), \ldots, u(\boldsymbol{x}_r)]^\top$, we can write these m constraints as:

$$W\mathbf{u}(X) > 0$$

Likelihood: $p(\mathcal{D}_m|\mathbf{u}(X)) = I_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X))$

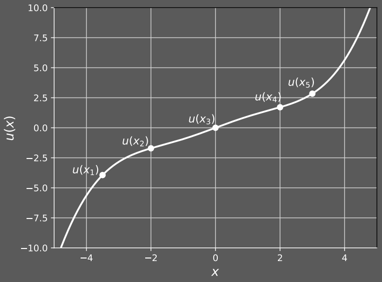
Prior: $p(\mathbf{u}(X)) = N(\mathbf{u}(X); \mathbf{0}, K_{\boldsymbol{\theta}}(X, X))$

Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = TN_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X); \mathbf{0}, K_{\boldsymbol{\theta}}(X, X))$

Posterior pred.: $p(\mathbf{u}(X^*)|\mathcal{D}_m) = \int p(\mathbf{u}(X^*)|\mathbf{u}(X))\, p(\mathbf{u}(X)|\mathcal{D}_m) d\mathbf{u}(X)$

What is a GP?

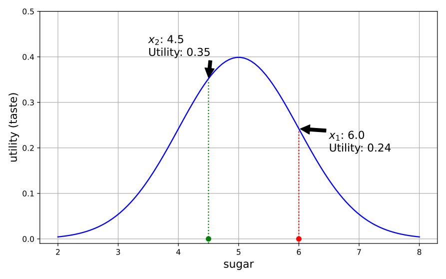
$$u \sim GP(m(x), k(x, x'))$$

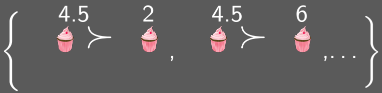


$$\begin{bmatrix} u(x_1) \\ u(x_2) \\ u(x_3) \\ u(x_4) \\ u(x_5) \end{bmatrix} \sim N\left(\begin{bmatrix} m(x_1) \\ m(x_2) \\ m(x_3) \\ m(x_4) \\ m(x_5) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_5) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_5) \\ \vdots & \vdots & \vdots & \vdots \\ k(x_5, x_1) & k(x_5, x_2) & \dots & k(x_5, x_5) \end{bmatrix} \right)$$

Typical Kernel: $k_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_{i=1}^{c} \frac{(x_i - x_i')^2}{2\ell_i^2} \right)$

Example

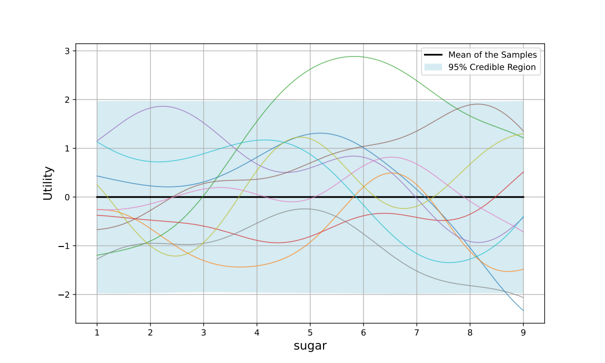




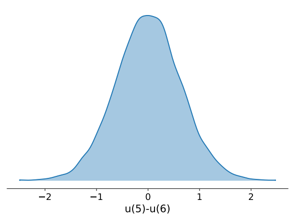
$\Rightarrow u(sugar)$ and predict $u(5) \overset{?}{>} u(6)$

Prior

$$\text{Prior:} \quad p(\mathbf{u}(x)) = GP(m(x), k(x, x'))$$

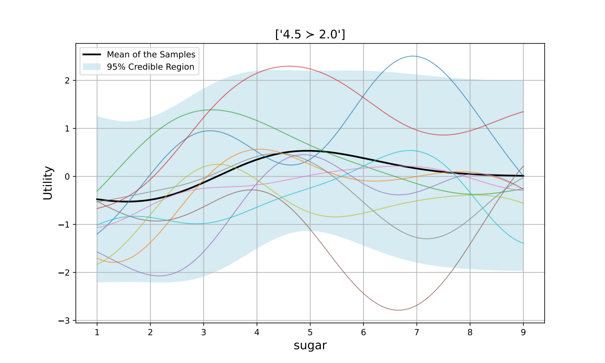


$$u(5) \overset{?}{>} u(6)$$

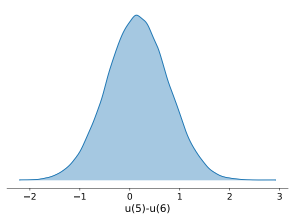


One preference

Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = TN_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X); \mathbf{0}, K_\theta(X, X))$

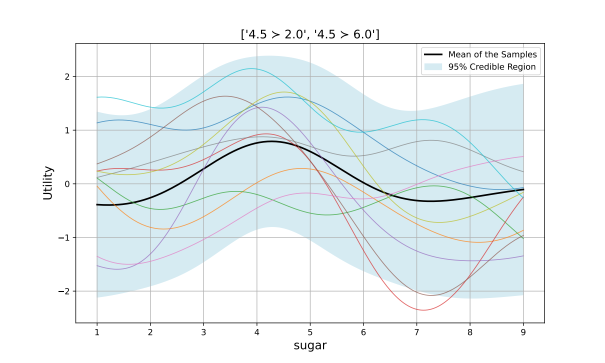


$u(5) \overset{?}{>} u(6)$

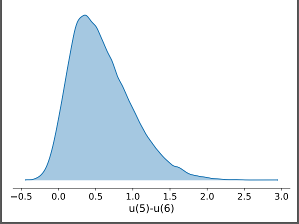


Two preferences

Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = TN_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X); \mathbf{0}, K_\theta(X,X))$

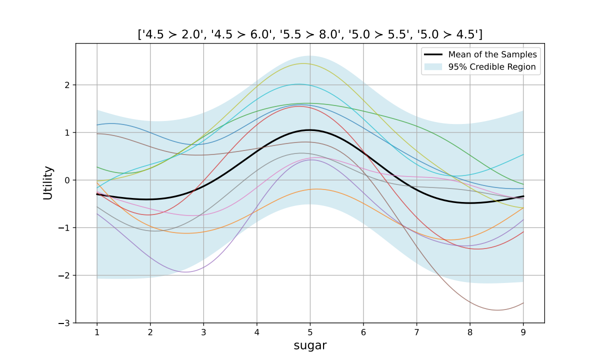


$u(5) \overset{?}{>} u(6)$

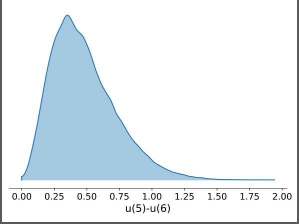


Five preferences

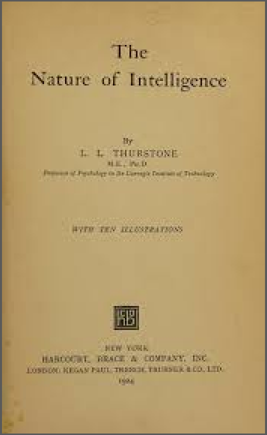
Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = TN_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X); \mathbf{0}, K_\theta(X,X))$

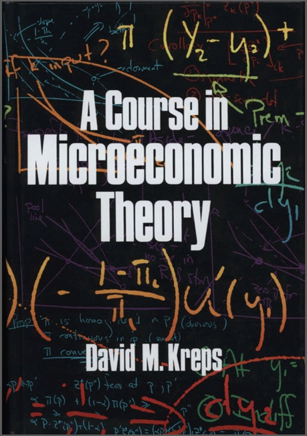


$u(5) \overset{?}{>} u(6)$



The Nature of Intelligence





Translating these ideas into statistical learning!

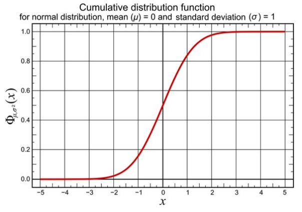
User's preferences may be inconsistent

1. Limit of discernability:

2. Noise in the observed utility

Robustify the model: Limit of discernability

$$\begin{cases} \mathbf{x}_i \succ \mathbf{x}_j & \text{with probability } \Phi\left(\frac{u(\mathbf{x}_i)-u(\mathbf{x}_j)}{\sigma}\right) \\ \mathbf{x}_j \succ \mathbf{x}_i & \text{with probability } 1-\Phi\left(\frac{u(\mathbf{x}_i)-u(\mathbf{x}_j)}{\sigma}\right) \end{cases}$$



Likelihood: $p(\mathcal{D}_m|\mathbf{u}(X)) = \prod_{s=1}^{m} \Phi\left(\frac{u(\mathbf{x}_l{}^{(s)})-u(\mathbf{x}_r{}^{(s)})}{\sigma}\right) = \Phi_m\left(\frac{W}{\sigma}\mathbf{u}(X)\right)$

Prior: $p(\mathbf{u}(X)) = N(\mathbf{u}(X); \mathbf{0}, K_\theta(X,X))$

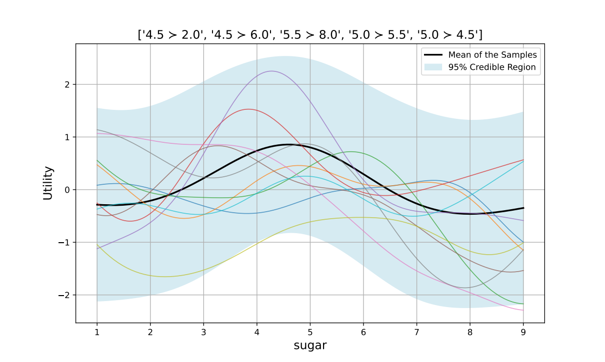
Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = \text{SkewNormal}(\mathbf{u}(X); \mathbf{0}, K_\theta(X,X), \Delta, \Gamma, \gamma)$

Posterior pred.: $p(\mathbf{u}(X^*)|\mathcal{D}_m) = \int p(\mathbf{u}(X^*)|\mathbf{u}(X))\, p(\mathbf{u}(X)|\mathcal{D}_m) d\mathbf{u}(X)$

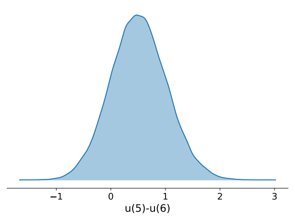
This model was derived in [Benavoli et al., 2021a, Benavoli et al., 2021b]. Originally proposed by [Chu and Ghahramani, 2005] using Laplace's approximation.

Five preferences

Posterior: $p(\mathbf{u}(X)|\mathcal{D}_m) = \text{SkewNormal}(\mathbf{u}(X); \mathbf{0}, K_{\theta}(X,X), \Delta, \Gamma, \gamma)$



$u(5) \overset{?}{>} u(6)$



User's preferences may be inconsistent due to time

- **Monday**:

 $\succ$ 

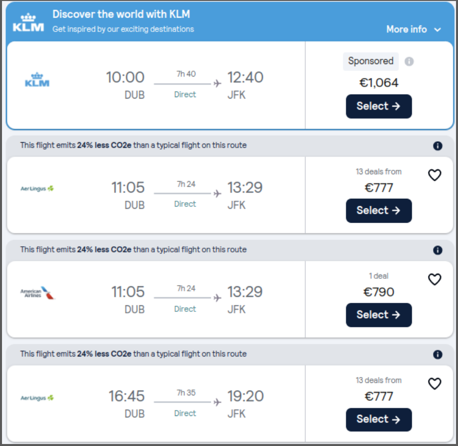
- **Tuesday**:

 $\succ$ 

Make the model to be aware of time:

$$\text{Prior:} \quad p(\mathbf{u}(X,t)) = N(\mathbf{u}(X,t); \mathbf{0}, K_{\boldsymbol{\theta}}((X,t),(X,t)))$$

More rationality



Objects are flights:

$$\mathbf{x}_i = \begin{bmatrix} \text{cost} \\ \text{duration} \\ \text{company} \\ \text{dep. time} \end{bmatrix}$$

Additional rationality constraints:

$$\begin{bmatrix} \text{cost} + \delta \\ \text{duration} \\ \text{company} \\ \text{dep. time} \end{bmatrix} \prec \begin{bmatrix} \text{cost} \\ \text{duration} \\ \text{company} \\ \text{dep. time} \end{bmatrix}$$

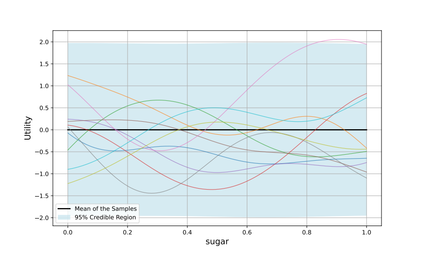
or, equivalently,

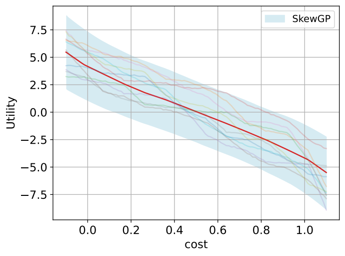
$$u(\text{cost} + \delta, \dots) < u(\text{cost}, \dots)$$

Similarly for *duration*.

Monotonic GP

$$u(cost) \sim GP(m(cost), k(cost, cost')) \quad \text{subject to} \quad \frac{d}{d\text{cost}} u < 0$$





When gambles are the objects

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \succ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Then there is a utility that represents the preference

$$u(\mathbf{x}_1) > u(\mathbf{x}_2)$$

Add monotonicity

$$\frac{\partial}{\partial x_{11}} u(\mathbf{x}_1) > 0, \quad \frac{\partial}{\partial x_{12}} u(\mathbf{x}_1) > 0$$

Given the preferences

$$\{\mathbf{x}_1 \succ \mathbf{x}_2, \quad \mathbf{x}_1 \succ \mathbf{x}_4 \ldots\} \quad \text{learn} \quad u$$

When gambles are the objects

$\succ$

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \succ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Then there is a utility that represents the preference

$$u(\mathbf{x}_1) > u(\mathbf{x}_2)$$

Add monotonicity

$$\frac{\partial}{\partial x_{11}} u(\mathbf{x}_1) > 0, \quad \frac{\partial}{\partial x_{12}} u(\mathbf{x}_1) > 0$$

Given the preferences

$$\{\mathbf{x}_1 \succ \mathbf{x}_2, \quad \mathbf{x}_1 \succ \mathbf{x}_4 \ldots\} \quad \text{learn} \quad u$$

Add linearity

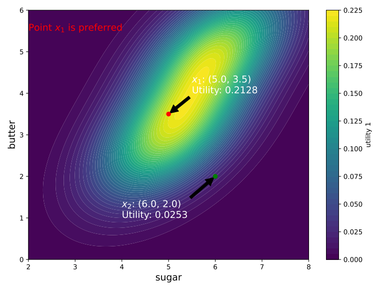
$$u(\mathbf{x}_1) = w_1 x_{11} + w_2 x_{12} \quad \text{equiv.} \quad u(\mathbf{x}) \sim GP(0, \mathbf{x}\mathbf{x}^\top)$$

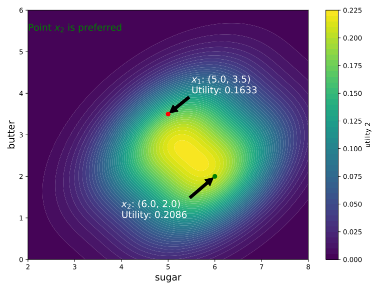
Multiple utilities and incompleteness

(**cpu**,drive)  ≻ (**cpu**,drive) 

(cpu,**drive**)  ≻ (cpu,**drive**) 

(**cpu**,drive)  ≻ (**cpu**,drive) 





Choice functions

$A = \{$  ,  ,  ,  $\}$, $C(A) = \{$  ,  $\}$

- rationality criteria for choices (Chernoff, Expansion, Aizerman, Sen-α..)

  $A = \{$  ,  ,  $\}$, $C(A) = \{$  ,  $\}$
- utility representation
  - A cupcake is selected if it is the best with respect to (at least) one utility (in IP e-admissibility)

Choice functions

$A = \{$  ,  ,  ,  $\}$, $\quad C(A) = \{$  ,  $\}$

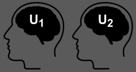
- rationality criteria for choices (Chernoff, Expansion, Aizerman, Sen-α..)

  $A = \{$  ,  ,  $\}$, $\quad C(A) = \{$  ,  $\}$

- utility representation
  - A cupcake is selected if it is the best with respect to (at least) one utility (in IP e-admissibility)

$$p(C(A), A|u_1(X), \ldots, u_L(X)) = \prod_{\{\mathbf{o},\mathbf{v}\} \in C_\sharp(A)} \left(1 - \prod_{i=1}^{L} I_{u_i(\mathbf{o})>u_i(\mathbf{v})}(\mathbf{u}(X)) - \prod_{i=1}^{L} I_{u_i(\mathbf{v})>u_i(\mathbf{o})}(\mathbf{u}(X))\right)$$

$$\prod_{\mathbf{v} \in R(A)} \left(\prod_{i=1}^{L} \left(1 - \prod_{\mathbf{o} \in C(A)} I_{u_i(\mathbf{v})>u_i(\mathbf{o})}(\mathbf{u}(X))\right)\right)$$



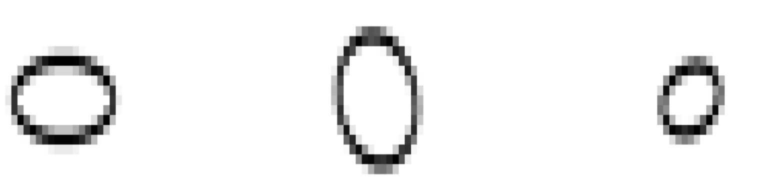
$u_1 \sim GP(0, k)$

$u_2 \sim GP(0, k)$

Example: choosing ellipses

Two criteria:

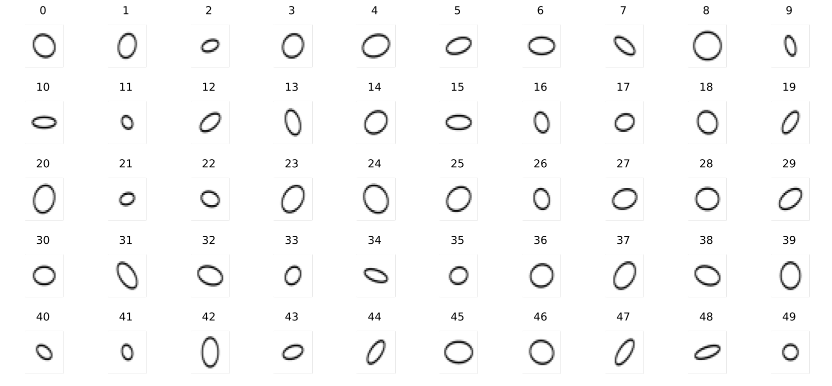
- "being closer to a circle"
- "being aligned to the axes"



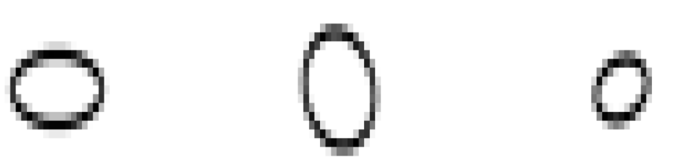
The student selected the left and right-ellipse.

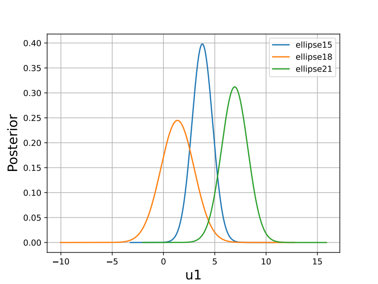
Example: choosing ellipses

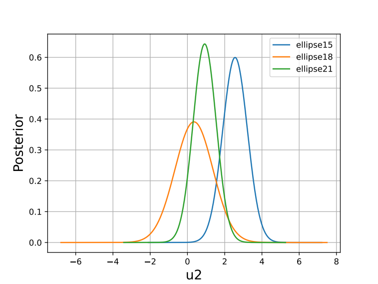
Dataset: 50 images 25×25 pixels and 160 choices



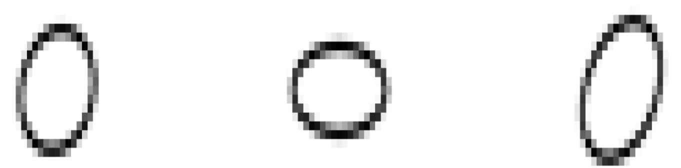
Example: choosing ellipses

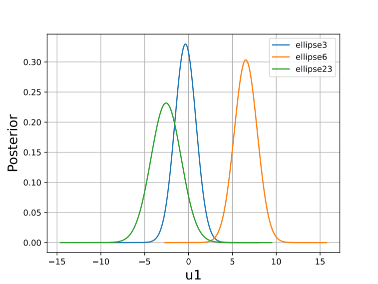


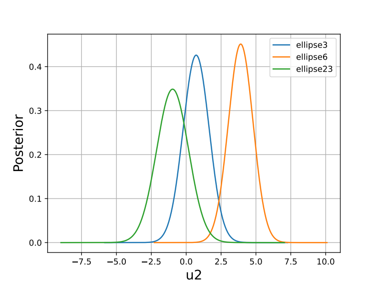




Example: choosing ellipses

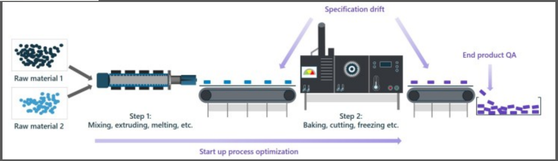




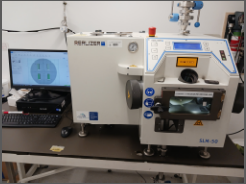
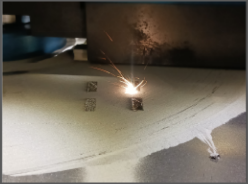
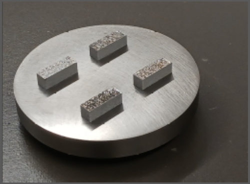


I could make ellipses that the model predicts you would like the most.

Smart manufacturing



Smart manufacturing

Block 0 $\mathbf{x}_0 = [\mathit{Power},\ \mathit{Speed},\ \mathit{HatchSpacing}] = [300, 100, 40]$

Block 1 $\mathbf{x}_1 = [90, 700, 140]$

Block 2 $\mathbf{x}_2 = [70, 250, 50]$

Block 3 $\mathbf{x}_3 = [40, 550, 110]$

Thank you

Preferences are everywhere!

desirability

$$u(x_i) > 0$$

preference

$$u(x_i) > u(x_j)$$

ranking

$$\tau_{2star} > u(x_i) > \tau_{1star}$$

interval

$$7 > u(x_i) > 6$$

precise measurements

$$6.3 + \epsilon > u(x_i) > 6.3$$

Thank you

Preferences are everywhere!

desirability

$u(x_i) > 0$

preference

$u(x_i) > u(x_j)$

ranking

$\tau_{2star} > u(x_i) > \tau_{1star}$

interval

$7 > u(x_i) > 6$

precise measurements

$6.3 + \epsilon > u(x_i) > 6.3$

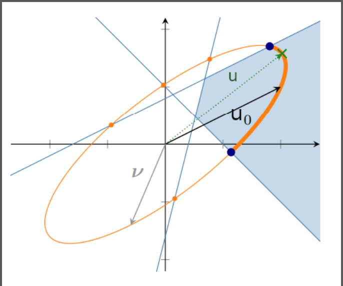
Linear elliptical slice sampling [Gessner et al., 2020]

1. Given previous sample $u_0(X)$
2. Sample $\nu(X) \sim N(\mathbf{u}(X); \mathbf{0}, K_{\boldsymbol{\theta}}(X, X))$
3. Define the ellipse

$$u_0(X)\cos(\phi) + \nu(X) sin(\phi), \quad \phi \in [0, 2\pi)$$

4. Intercept ellipse with $\{W\mathbf{u}(X) > 0\}$
5. Sample ϕ at random inside the intersection

$$u_1(X) = u_0(X)\cos(\phi) + \nu(X) sin(\phi)$$



Learning Hyperparameters

Marginal likelihood:

$$\arg\max_{\boldsymbol{\theta}} p(\mathcal{D}_m|\boldsymbol{\theta}) = \int I_{\{W\mathbf{u}(X)>0\}}(\mathbf{u}(X))N(\mathbf{u}(X);\mathbf{0},K_{\boldsymbol{\theta}}(X,X))d\mathbf{u}(X)$$

Approximation:

1. Approximation of numerical integration for Truncated Gaussian;
2. Laplace's approximation;
3. Variational approximation.

References I

Benavoli, A., Azzimonti, D., and Piga, D. (2021a).
Preferential bayesian optimisation with skew gaussian processes.
In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 1842–1850.

Benavoli, A., Azzimonti, D., and Piga, D. (2021b).
A unified framework for closed-form nonparametric regression, classification, preference and mixed problems with skew gaussian processes.
Machine Learning, 110(11):3095–3133.

References II

Chu, W. and Ghahramani, Z. (2005).
Preference learning with gaussian processes.
New York, NY, USA. Association for Computing Machinery.

Dekel, O., Singer, Y., and Manning, C. D. (2003).
Log-linear models for label ranking.
Advances in neural information processing systems, 16.

Gessner, A., Kanjilal, O., and Hennig, P. (2020).
Integrals over gaussians under linear domain constraints.
In *International Conference on Artificial Intelligence and Statistics*, pages 2764–2774. PMLR.

References III

Har-Peled, S., Roth, D., and Zimak, D. (2002).
Constraint classification for multiclass classification and ranking.
Advances in neural information processing systems, 15.