Introduction

High level Concept Explanation

Problem Enviornment

An individual of type $X \in \mathcal{X}$ arrives independently according to probability distribution $\Gamma_X \in \Delta(\mathcal{X})$. We will respond with an action $Y \in \mathcal{Y}$ (which can be random) that then earns us a utility $U(X,Y) \in \mathbb{R}$. Once X and Y are realized the value of $U(X,Y)$ is drawn independently from $\Gamma_U(X,Y) \in \Delta(\mathbb{R})$. In this way Γ_U is a function on $\mathcal{X} \times \mathcal{Y}$ which determines the irreducible uncertainty in the outcomes of our problem. Our goal is to maximize U.

Decision Rule

To determine Y a decision rule $D : \mathcal{X} \to \Delta(\mathcal{Y})$ is implemented. A decision rule captures our system's response to the problem. As a utility-maximizer we hope to take an action from,

$$\mathcal{Y}_X^* = \arg\max_Y \mathbb{E}[U(X,Y)|X].$$

Thus, when evaluating a decision rule, we can focus on minimizing the deviations taken from this optimal set. This gives us the expected loss function for a decision rule D,

$$L(D) = \mathbb{E}[U(X, Y^*(X)) - U(X, D(X))],$$

where $Y^*(X) \in \mathcal{Y}_X^*$. Regardless of whether we have direct control over the formation of D, we want to take actions to minimize $L(D)$.

Decision-Maker

The decision-maker is the creator of the decision rule. They earn a utility of,

$$V(X,Y) = U(X,Y) + \delta(X,Y).$$

In this way δ exactly encodes the deviations between our utility and the decision-maker's utility. The problem type X is not known to the decision-maker. Instead the problem realization induces a probability distribution over which the decision-maker assume's X was drawn. We call this $\Phi : X \to \Delta(\mathcal{X})$, the decision-maker's interpretation of the problem. Together these make up the type of the decision-maker $\theta = (\Phi, \delta)$. If $\delta = 0$, we say the decision-maker is preference aligned (as the decision maker acts exactly as we would want them to subject to their understanding of the problem).

Assistive Signal