CS 497 SML
Spring 2003
Scribe Notes

by Shai Sachs

28 Jan. 2003

Agenda

(Today's topics in bold)

- Decision Theory
  - Nondeterministic decisions
  - Bayesian decision theory
  - **Utility Theory**
  - **Criticisms**
  - Multi-objective optimality
- Sequential decision making
  - States
  - Stage-additive loss

Overview

An overview of the field of decision theory and its subfields is pictured in Figure 1.

The basic task of classification can be described as follows. We are given y, a *feature vector*, where $y \in Y$, and Y is the set of all possible feature vectors. The set of possible classes is Ω. Given an object with a feature vector y, we wish to determine the correct class $\omega \in \Omega$ of the object.

Ideally, we are given $P(y|\omega)$ and $P(\omega)$, the *prior distribution* over the classes. The probability $P(\omega)$ gives the probability that an object falls in the class ω.

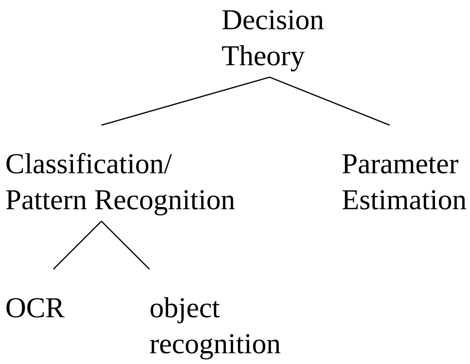

Figure 1: An overview of decision theory.

1 Parameter Estimation

One subfield of Decision Theory is parameter estimation. The goal is to estimate the value of some parameter, given some observation of the parameter. We consider the parameter to be some fixed constant, and denote the set of all possible parameters (the *parameter space*) as X.

Using our notation from decision theory, we have $\Theta = X$. The parameter we are trying to estimate is nature's choice.

Since the goal is the guess the correct value of the parameter, the set of actions U is also equal to X; that is, the human player's action is to choose some valid value of the parameter as her guess. Therefore, we have $X = \Theta = U$.

Further, we have an observation about the parameter, y; we denote the set of all possible observations by Y. Clearly, $X \subseteq Y$.

Suppose we have $X = [0, 1] \subseteq \mathbb{R}$, $p(y|x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-y)^2}{2\sigma^2}\right)$, and $p(x) = 1$. We interpret these probability density functions as follows: $p(y|x)$ tells us that there is some Gaussian noise in the observation (that is, our observations of the parameters, over many trials, will be concentrated around the true parameter in a Gaussian distribution); further, $p(x)$ tells us that each parameter is equally likely.

Finally, we choose a loss function which measures the estimation error (that is, the difference between our estimate and the true parameter). We use $L(u, x) = (u - x)^2$.

We wish to choose the input u which minimizes our risk; we therefore choose u which minimizes

$$R(u) = \int_x L(u, x) p(y|x) p(x) dx \tag{1}$$

Note that when Equation 1 is multiplied by $\frac{1}{p(y)}$, by Bayes' rule we have

$$R(u) = \int_x L(u,x)p(y|x)p(x)\frac{1}{p(y)}dx = \int_x L(u,x)p(x|y)dx$$

Then the expression for $R(u)$ becomes exactly analogous to the discrete form of the risk function from our previous lecture. Since $p(y)$ is constant over X, we may remove it from the integral without affecting the correct choice of u. Therefore $R(u)$ in Equation 1 is not exactly the risk function, but it is closely related.

2 Utility Theory

Utility theory asks: Where does the loss function L come from? In other words, how useful is one loss compared against another?

In utility theory we replace "loss" with "reward"; the human wants to maximize reward. Note that this convention can easily be inverted to return the the usual "loss" convention in decision theory.

2.1 Choosing a Good Reward

Let u_1 be some fairly unpleasant task (such as writing scribe notes), and u_2 be the act of doing nothing. We consider the problem of choosing a good reward function using the following examples.

1. Let $R(u_1) = 1000$, and $R(u_2) = 0$. For a poor graduate student, it may be worthwhile to write scribe notes for \$ 1000, so the student will probably do u_1 in this scenario. One difficulty in this scenario is that we haven't considered the possible cost to the human of each action.

2. Let $R(u_1) = 10001000$, and $R(u_2) = 10000000$. Although the relative reward is the same, the action chosen is probably different! This is so because the value (or utility) of money decreases as we have more of it.

3. Let $R(u_1) = 10000$ and $R(u_2) = 25,000$ with probability $\frac{1}{2}$, and 0 with probability $\frac{1}{2}$. In this scenario some conservative students may choose u_1, to guarantee a reward; while more adventurous gamblers may choose u_2, as the expected gain is greater.

4. Let $R(u_1) = 100$, and $R(u_2) = 250$ with probability $\frac{1}{2}$, and 0 otherwise; allow the student to choose an action (and collect the corresponding reward) 100 times. The expected reward for each action remains the same, but we imagine more students will choose u_2 100 times than those that will choose u_1 100 times. This is so because "repeatability" is important in games of expectation; that is, the true outcome of a game is more likely nearer the expected value if the number of trials increases.

	u_1	u_2
θ_1	1	0
θ_2	3	2

Figure 2:

The goal of utility theory is to construct reward functions that give the right expected value for a game, given the preferences of the human player.

We call the set of all possible rewards for a given game the *reward space*, and denote it by $\mathcal{R}$. For example, in scenario 3, we have $\mathcal{R} = \{0, 10000, 25000\}$.

Consider game with nature depicted in Figure 2. Suppose $P(\theta_1) = \frac{1}{4}$ and $P(\theta_2) = \frac{3}{4}$. Then choosing u_1 implies we get $R(u_1) = 1$ with probability $\frac{1}{4}$, and $R(u_1) = 3$ with probability $\frac{3}{4}$. We may consider the prior distribution over Θ as giving us a probability distribution over $\mathcal{R}$.

In general, we let $\mathcal{P}$ be the set of all probability distributions over $\mathcal{R}$, and let $P \in \mathcal{P}$ be one such distribution. We expect the human player to express her preference between any two such probability distributions, and we denote these preferences with the usual inequality operations. Thus $P_1 \leq P_2$ indicates that the human prefers P_1 no more than P_2, $P_1 = P_2$ indicates that the human has no preference among P_1, P_2, and so on.

We may then express the goal of utility theory as follows: we wish to find some function $V : \mathcal{R} \mapsto \mathbb{R}$ such that $P_1 < P_2$ iff $E^{P_1}[V(r)] < E^{P_2}[V(r)]$. That is, the expected value of a reward is greater under more preferred distributions over the reward space.

Note that computing V is difficult. However, we know (but will not prove) that V exists when the human is rational – that is, when her choices obey the Axioms of Rationality.

2.2 Axioms of Rationality

We say that a human is *rational* when the preferences she expresses among probability distributions over $\mathcal{R}$ obey the following axioms.

1. If $P_1, P_2 \in \mathcal{P}$, then either $P_1 \leq P_2$ or $P_2 \leq P_1$.
2. If $P_1 \leq P_2$ and $P_2 \leq P_3$ then $P_1 \leq P_3$.
3. If $P_1 < P_2$ then $\alpha P_1 + (1-\alpha)P_3 < \alpha P_2 + (1-\alpha)P_3$, for all $P_3 \in \mathcal{P}$ and $\alpha \in (0,1)$.

   This axiom is strange, but it merely means that no matter how much we "blend" P_1 and P_2 with some other distribution P_3, we will still prefer the "blended" P_2 to the "blended" P_1.
4. If $P_1 < P_2 < P_3$ then $\exists \alpha \in (0,1), \beta \in (0,1)$ such that $\alpha P_1 + (1-\alpha)P_3 < P_2$ and $P_2 < \beta P_1 + (1-\beta)P_3$.

This axiom means that no matter how good P_3 is, we can always blend a bit of it with P_1 to get a distribution less preferable than P_2; similarly, no matter how bad P_1 is, we can always blend a bit of it with P_3 to get a distribution more preferable than P_2.

3 Criticisms

We consider a few criticisms of decision theory:

1. The values of rewards are subjective. If they are provided by the human, then the process of making a decision may amount to "garbage in, garbage out."

2. It is difficult to assign losses.

3. Badly chosen loss functions can lead to bad decisions.

   One response to this criticism is *sensitivity analysis*, which claims that the decisions are not hypersensitive to the loss functions. Of course, if this argument is taken to far, then the value of decision theory itself is thrown into question.