increases. The graph below is a sample demand curve, where the demand schedule for the quantity of toilet paper demanded is graphed.

From this graph we can determine how many rolls of toilet paper will be purchased at what price. As can be seen from looking at this graph, it is negatively sloping. As one variable gets larger the other will become smaller, or when the price drops more is purchased. The whole demand curve "theory" is based on human behavior. It is logical to say that people will purchase more of a product when the price is cheaper.

In reality, if the price of a good rises the income (or assets) of the consumer will decrease. The people would not be able to buy the same goods as before because they cost more. Consumers can do two things; if the good is a normal good (previously defined), they would buy less of the good; if the good is an inferior good, they would buy more of the good. Thus, the income effect can be defined in this statement: When the price of a good falls, the expected outcome would be that the consumers would buy more because they have the money and can afford to buy more. The slope of the demand curve can be explained in terms of the income and substitution effects.

The quantity demanded of a good usually is a strong function of its price. Suppose an experiment is run to determine the quantity demanded of a particular product at different price levels, holding everything else constant. Presenting the data in tabular form would result in a demand schedule, an example of which is shown below.

Demand Schedule

Price Demanded	Quantity
5	10
4	17
3	26
2	38
1	53

The demand curve for this example is obtained by plotting the data:

Demand Curve

By convention, the demand curve displays quantity demanded as the independent variable (the x axis) and price as the dependent variable (the y axis). The law of demand states that quantity demanded moves in the opposite direction of price (all other things held constant), and this effect is observed in the downward slope of the demand curve. For basic analysis, the demand curve often is approximated as a straight line. A demand function can be written to

2. Income: an increase or decrease of consumer income will affect their disposable income and discretionary spending trends- increasing or decreasing demand
3. Population: the population of an area will affect demand. A larger population means more consumers and greater demand and vice a versa.
4. Income distribution: an even distribution of income will mean an increase for demand of luxury goods by low and middle income groups whereas an uneven distribution would lead to increased demand for products.

2. On the basis of the analysis of the case above, what is your opinion about legalizing marijuana in Canada?

There is no doubt that legalizing marijuana has its advantages, mainly in terms of increased revenues to the government due to corresponding increase in tax collections. In also has other concomitant benefits like reducing government spend on law enforcement related to controlling proliferation of narcotics. A legalized regime would also ensure greater control on the market demand and the consequent supply of marijuana. However, there are far greater implications due to legalization of marijuana. This includes increased incidence of ill health amongst the population and could increase black marketing in case the prices of marijuana are kept at a very high level. Considering the implications on the health of the general population, it would not be correct to legalize marijuana purchase and sale in Canada.

Case let 2

Companies that attend to productivity and growth simultaneously manage cost reductions very differently from companies that focus on cost cutting alone and they drive growth very differently from companies that are obsessed with growth alone. It is the ability to cook sweet and sour that under grids the remarkable performance of companies likes Intel, GE, ABB and Canon. In the slow growth electro-technical business, ABB has doubled its revenues from $17 billion to $35 billion, largely by exploiting new opportunities in emerging markets. For example, it has built up a 46,000 employee organization in the Asia Pacific region, almost from scratch. But it has also reduced employment in North America and Western Europe by 54,000 people. It is the hard squeeze in the north and the west that generated the resources to support ABB's massive investments in the east and the south. Everyone knows about the staggering ambition of the Ambanis, which has fuelled Reliance's evolution into the largest private company in India. Reliance has built its spectacular rise on a similar ability to cook sweet and sour. What people may not be equally familiar with is the relentless focus on cost reduction and productivity growth that pervades the company. Reliance's employee cost is 4 per cent of revenues, against 15-20 per cent of

material expensive for our local industry, when on the other hand the price of end products were falling. Hence the industry got squeezed both ways. To top the agonies of the industry, in February 2009 the government introduced with retrospective effect five percent license against exports of cotton, hence making our cotton cheaper for our competing nations – We need to rethink whether we want value addition on Indian cotton. We have been talking about a 'fiber policy' for the last two years, but still there is nothing offered except promises.

On the manmade fiber side, due to our inherently high raw material prices and import duties, we are clearly out priced. Despite import duties, many manmade textile intermediaries are being imported into the country.

Secondly, let us analyze the technology access and its cost. India has access to the best global technology in addition to its locally available machinery. However, apart from ginning and partly spinning, we are largely dependent on imported technology. Europe is considered a hub for it, but the severe appreciation of its currency has made it more expensive by about 20 percent over the last few years (European machinery is anyways not cheap), which is further compounded by import duty despite export obligations. The Chinese textile machinery industry is much more developed and, in fact, is selling all over the world including India. This, coupled with rising interest rates over the last two years has made capital expensive and projects like spinning mills difficult. Despite the Textile Upgradation Fund, the effective rates are about 7-8 percent which is high compared to today's international funding rates. Even our discounted working capital rate for exporters at 7 percent means a labor plus five percent rate, which makes us uncompetitive. Due to the recession interest rates have crashed globally; but India rates have only fallen marginally and due to the risk premium going up the positive impact is further reduced.

Thirdly, we can talk about labor – our favorite listing when it comes to analysis of our advantages. However, in absolute terms we are higher than many of our competitors like Bangladesh, Pakistan, and other developing countries and if adjusted for productivity we are higher than all major textile nations like China, Indonesia, Thailand, and Vietnam.

Further, we have a restriction on labor flexibility, making it very difficult for the garment industry, which has seasonal loads of work – thereby allowing very limited space for companies to adjust to the changing times and seasonal demands.

The Government has a minimum guarantee employment program of 100 days (NREGS). Why can't it,on a similar basis, allow a 200 hundred days guaranteed employment to laborers of the garment industry? The important point is that the industry would pay higher wages and would also enhance skills of the workers.

country's regulations, or that they achieve the importing country's appropriate level of protection13.

The concept of risk assessment is an environmental concept which is used in risk management and is relevant to the precautionary approach. It has been incorporated into the SPS Agreement as a means for determining the level of sanitary or phytosanitary protection which is appropriate for each country14. SPS Article 5 invites Members to ensure that their measures "are based on an assessment ... of the risks to human, animal or plant life or health ..", taking into account, on the one hand, "available scientific evidence; relevant processes and production methods; relevant inspection, sampling and testing methods; prevalence of specific diseases or pests; existence of pest- or disease- free areas; relevant ecological and environmental conditions; and quarantine or other treatment"; and on the other hand relevant economic factors, such as "the potential damage in terms of loss of production or sales in the event of the entry, establishment or spread of pest or disease; the costs of control or eradication in the territory of the importing Member; and the relative cost-effectiveness of alternative approaches to limiting risks".

2. What role does a decision tree play in business decision-making? Illustrate the choice between two investment projects with the help of decision tree assuming hypothetical conditions about the states of nature, probability distribution and corresponding pay-offs.

Answer

Introduction

In many problems chance (or probability) plays an important role. Decision analysis is the general name that is given to techniques for analysing problems containing risk/uncertainty/probabilities. Decision trees are one specific decision analysis technique and we will illustrate the technique by use of an example.

Example

A company faces a decision with respect to a product (codenamed M997) developed by one of its research laboratories. It has to decide whether to proceed to test market M997 or whether to drop it completely. It is estimated that test marketing will cost £100K. Past experience indicates that only 30% of products are successful in test market.

If M997 is successful at the test market stage then the company faces a further decision relating to the size of plant to set up to produce M997. A small plant will cost £150K to build and produce 2000 units a year whilst a large plant will cost £250K to build but produce 4000 units a year.

or

If the parents are not visiting and it is rainy, then stay in.

Of course, this is just a re-statement of the original mental decision making process we described. Remember, however, that we will be programming an agent to learn decision trees from example, so this kind of situation will not occur as we will start with only example situations. It will therefore be important for us to be able to read the decision tree the agent suggests.

Decision trees don't have to be representations of decision making processes, and they can equally apply to categorisation problems. If we phrase the above question slightly differently, we can see this: instead of saying that we wish to represent a decision process for what to do on a weekend, we could ask what kind of weekend this is: is it a weekend where we play tennis, or one where we go shopping, or one where we see a film, or one where we stay in? For another example, we can refer back to the animals example from the last lecture: in that case, we wanted to categorise what class an animal was (mammal, fish, reptile, bird) using physical attributes (whether it lays eggs, number of legs, etc.). This could easily be phrased as a question of learning a decision tree to decide which category a given animal is in, e.g., if it lays eggs and is homeothermic, then it's a bird, and so on.

Learning Decision Trees Using ID3

Specifying the Problem

We now need to look at how you mentally constructed your decision tree when deciding what to do at the weekend. One way would be to use some background information as axioms and deduce what to do. For example, you might know that your parents really like going to the cinema, and that your parents are in town, so therefore (using something like Modus Ponens) you would decide to go to the cinema.

Another way in which you might have made up your mind was by generalising from previous experiences. Imagine that you remembered all the times when you had a really good weekend. A few weeks back, it was sunny and your parents were not visiting, you played tennis and it was good. A month ago, it was raining and you were penniless, but a trip to the cinema cheered you up. And so on. This information could have guided your decision making, and if this was the case, you would have used an inductive, rather than deductive, method to construct your decision tree. In reality, it's likely that humans reason to solve decisions using both inductive and deductive processes.

We can state the problem of learning decision trees as follows:

We have a set of examples correctly categorised into categories (decisions). We also have a set of attributes describing the examples, and each attribute has a finite set of values which it can possibly take. We want to use the examples to learn the structure of a decision tree which can be used to decide the category of an unseen example.

Assuming that there are no inconsistencies in the data (when two examples have exactly the same values for the attributes, but are categorised differently), it is obvious that we can always construct a decision tree to correctly decide for the training cases with 100% accuracy. All we have to do is make sure every situation is catered for down some branch of the decision tree. Of course, 100% accuracy may indicate overfitting.

The basic idea

In the decision tree above, it is significant that the "parents visiting" node came at the top of the tree. We don't know exactly the reason for this, as we didn't see the example weekends from which the tree was produced. However, it is likely that the number of weekends the parents visited was relatively high, and every weekend they did visit, there was a trip to the cinema. Suppose, for example, the parents have visited every fortnight for a year, and on each occasion the family visited the cinema. This means that there is no evidence in favour of doing anything other than watching a film when the parents visit. Given that we are learning rules from examples, this means that if the parents visit, the decision is already made. Hence we can put this at the top of the decision tree, and disregard all the examples where the parents visited when constructing the rest of the tree. Not having to worry about a set of examples will make the construction job easier.

This kind of thinking underlies the ID3 algorithm for learning decisions trees, which we will describe more formally below. However, the reasoning is a little more subtle, as (in our example) it would also take into account the examples when the parents did not visit.

Entropy

Putting together a decision tree is all a matter of choosing which attribute to test at each node in the tree. We shall define a measure called information gain which will be used to decide which attribute to test at each node. Information gain is itself calculated using a measure called entropy, which we first define for the case of a binary decision problem and then define for the general case.

Given a binary categorisation, C, and a set of examples, S, for which the proportion of examples categorised as positive by C is p_+ and the proportion of examples categorised as negative by C is p_-, then the entropy of S is:

$$Entropy(S) = -p_+ \, log_2(p_+) - p_- \, log_2(p_-)$$

The reason we defined entropy first for a binary decision problem is because it is easier to get an impression of what it is trying to calculate. Tom Mitchell puts this quite well:

"In order to define information gain precisely, we begin by defining a measure commonly used in information theory, called entropy that characterizes the (im)purity of an arbitrary collection of examples."

Imagine having a set of boxes with some balls in. If all the balls were in a single box, then this would be nicely ordered, and it would be extremely easy to find a particular ball. If, however, the balls were distributed amongst the boxes, this would not be so nicely ordered, and it might take quite a while to find a particular ball. If we were going to define a measure based on this notion of purity, we would want to be able to calculate a value for each box based on the number of balls in it, then take the sum of these as the overall measure. We would want to reward two situations: nearly empty boxes (very neat), and boxes with nearly all the balls in (also very neat). This is the basis for the general entropy measure, which is defined as follows:

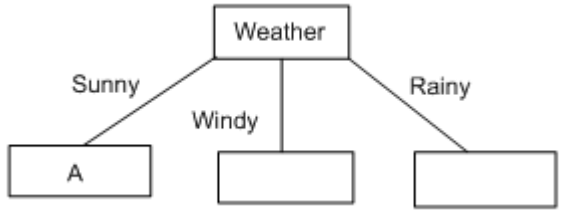
Given an arbitrary categorisation, C into categories c_1, ..., c_n, and a set of examples, S, for which the proportion of examples in c_i is p_i, then the entropy of S is:

$$Entropy(S) = \sum_{i=1}^{n} -p_i \log_2(p_i)$$

This measure satisfies our criteria, because of the -p*log2(p) construction: when p gets close to zero (i.e., the category has only a few examples in it), then the log(p) becomes a big negative number, but the p part dominates the calculation, so the entropy works out to be nearly zero. Remembering that entropy calculates the disorder in the data, this low score is good, as it reflects our desire to reward categories with few examples in. Similarly, if p gets close to 1 (i.e., the category has most of the examples in), then the log(p) part gets very close to zero, and it is this which dominates the calculation, so the overall value gets close to zero. Hence we see that both when the category is nearly - or completely - empty, or when the category nearly contains - or completely contains - all the examples, the score for the category gets close to zero, which models what we wanted it to. Note that 0*ln(0) is taken to be zero by convention.

Information Gain

We now return to the problem of trying to determine the best attribute to choose for a particular node in a tree. The following measure calculates a numerical value for a given attribute, A, with respect to a set of examples, S. Note that the values of attribute A will range



Now we have to fill in the choice of attribute A, which we know cannot be weather, because we've already removed that from the list of attributes to use. So, we need to calculate the values for Gain(S_{sunny}, parents) and Gain(S_{sunny}, money). Firstly, Entropy(S_{sunny}) = 0.918. Next, we set S to be S_{sunny} = {W1,W2,W10} (and, for this part of the branch, we will ignore all the other examples). In effect, we are interested only in this part of the table:

Weekend (Example)	Weather	Parents	Money	Decision (Category)
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W10	Sunny	No	Rich	Tennis

Hence we can calculate:

Gain(S_{sunny}, parents) = 0.918 - (|S_{yes}|/|S|)*Entropy(S_{yes}) - (|S_{no}|/|S|)*Entropy(S_{no})
= 0.918 - (1/3)*0 - (2/3)*0 = 0.918

Gain(S_{sunny}, money) = 0.918 - (|S_{rich}|/|S|)*Entropy(S_{rich}) - (|S_{poor}|/|S|)*Entropy(S_{poor})
= 0.918 - (3/3)*0.918 - (0/3)*0 = 0.918 - 0.918 = 0

Notice that Entropy(S_{yes}) and Entropy(S_{no}) were both zero, because S_{yes} contains examples which are all in the same category (cinema), and S_{no} similarly contains examples which are all in the same category (tennis). This should make it more obvious why we use information gain to choose attributes to put in nodes.

Given our calculations, attribute A should be taken as parents. The two values from parents are yes and no, and we will draw a branch from the node for each of these. Remembering that we replaced the set S by the set S_{Sunny}, looking at S_{yes}, we see that the only example of this is W1. Hence, the branch for yes stops at a categorisation leaf, with the category being Cinema. Also, S_{no} contains W2 and W10, but these are in the same category (Tennis). Hence the branch for no ends here at a categorisation leaf. Hence our upgraded tree looks like this: