Section Notes 8
Markov Chains and MDPs

Applied Math / Engineering Sciences 121

Week of November 18, 2019

Goals for the week

- be able to define stochastic processes and identify the Markovian Property.
- understand the properties of states and classes.
- be comfortable modeling dynamic systems with Markov chains and solving for the steady state (for ergodic chains)
- be familiar with Markov Decision Processes and different ways to determine the best policy.

Contents

1 Markov Chains

Markov Chains are incredibly useful constructs for modeling dynamic processes. Let's go over the related definitions, properties, and uses.

1.1 Definition

Exercise 1

What is a Stochastic Process?

End Exercise 1

Exercise 2

If a Stochastic Process has the **Markovian Property**, what does that mean?

End Exercise 2

Exercise 3

If a Stochastic Process with Markovian Property is also **Stationary**, what does that mean?

End Exercise 3

Exercise 4

If a Stochastic Process is both **Stationary** and **Markovian**, then what do we call it?

End Exercise 4

A Markov Chain can be described by a set of states $S = \{0, ..., m-1\}$, and a transition matrix whose entries p_{ij} define the probability of moving from state i into state j in one step:

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0\ m-1} \\ p_{10} & p_{11} & \cdots & p_{1\ m-1} \\ \cdots & \cdots & & \cdots \\ p_{m-1\ 0} & p_{m-1\ 1} & \cdots & p_{m-1\ m-1} \end{bmatrix}$$

Using this matrix we can define the probability of moving from state i to state j after t steps: $\mathbf{P}^{(t)}$. Below we denote a single entry in this latter matrix by $p_{ij}^{(t)}$.

1.2 Properties

Fill in the blank:

Exercise 5

State j is ____________________ from state i if $\exists t \geq 0 : p_{ij}^{(t)} > 0$

End Exercise 5

Fill in the second blank below with the previous answer and then fill in the first blank:

Exercise 6

States i and j are said to ____________________ if state j is ____________________ from i and visa versa.

End Exercise 6

This property we defined in exercise 6:

- Holds between any state and itself
- Is transitive
- Partitions the state into **equivalence classes**, where the states in each class all share the property with each other.

Exercise 7

A **Markov Chain** is ____________________ if and only if it has one such **class**.

End Exercise 7

Exercise 8

If, after an infinite amount of time, the process will definitely return to a state, that state is called ____________________. Otherwise it is called ____________________.

End Exercise 8

All states in a **Class** share this property as well.

Exercise 9

A state is ____________ if the process will never leave that state once entering it.

End Exercise 9

Exercise 10

The ____________ of a state i is the largest integer $d(i) \geq 1$ such that the chain can return to i only at multiples of $d(i)$.

End Exercise 10

Exercise 11

A state i is called ____________ if the period is $d(i) = 1$.

End Exercise 11

Fill in the *second* blank below with the previous answer and then fill in the first blank.

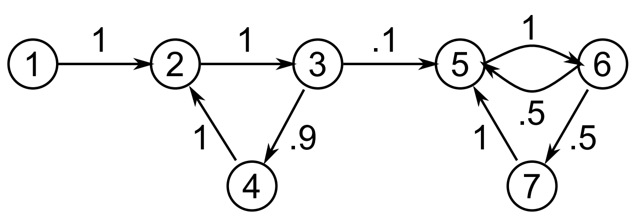
Exercise 12

An irreducible **Markov Chain** is ____________ if all of its states are ____________.

End Exercise 12

1.3 Example

Consider the following **Markov Chain**:



Exercise 13

Do states 1 and 2 communicate? Why or why not?

End Exercise 13

Exercise 14

What are the classes in this Markov Chain?

End Exercise 14

Exercise 15

Which states are *Recurrent*?

End Exercise 15

Exercise 16

Is this chain *Irreducible*?

End Exercise 16

Exercise 17

Are there any absorbing states?

End Exercise 17

Exercise 18

What is the period of each state?

End Exercise 18

Exercise 19

Is this Markov Chain ergodic?

End Exercise 19

By completing the above exercises, you should now:

- be comfortable with the definition and properties of Markov chains
- be able to determine the properties of the states and chain.

1.4 Steady-State Probabilities

For any *ergodic* **Markov Chain**

$$\lim_{t \to \infty} p_{ij}^{(t)} = \pi_j > 0$$

exists and is independent of i. The π_j represents the probability of the process being state j after a large number of transitions. $\pi = (\pi_0, \ldots, \pi_{m-1})$ is called the **stationary distribution**. We can specify a set of balance equations to determine these **steady-state probabilities**:

$$\pi_j = \sum_i \pi_i p_{ij} \ \forall \ j = 0, ..., m-1$$

$$\sum_j \pi_j = 1$$

Equations can be solved by standard linear algebra techniques (Gaussian elimination, finding the inverse, etc.)

1.5 Modeling

Many dynamic systems can be usefully modeled as Markov Chains. Let's try one:

Exercise 20

Suppose there are only four items on TV: Good Shows, Bad Shows, Good Show Commercials and Bad Show Commercials. Write down a **Markov Chain** that describes the dynamics of what is available on TV. Make up your probabilities.

End Exercise 20

Exercise 21

Now solve for the steady state of the Markov Chain (use the one provided on the board, rather than your own). What does this imply about TV?

End Exercise 21

By completing the above exercises, you should now:

- be able to model a Markov chain (determine states and transition probabilities)
- be comfortable solving for the steady state of an ergodic Markov chain

2 Markov Decision Processes

A Markov Decision Process has four parts:

- A set of states S
- A set of possible actions A
- A reward function R(s,a)
- A transition probability function $P(s, a, s') \in [0, 1]$

The Markov Property is implicit in this definition of the transition function: The effects of an action in a state are independent of any prior states.

This describes the dynamics of the system. In many cases, our goal is to find a **policy** given these dynamics. A deterministic policy $\mu : S \to A$ is a mapping from states to actions, i.e. what you should do in state S. Similarly we can define a stochastic policy that maps a state to a distribution of possible actions.

Exercise 22

Given an MDP and a policy, what should you do? More generally, what are the dynamics of the resultant system?

End Exercise 22

2.1 Objective Functions

An objective function defines which of the many possible **policies** is best. One can simply use a finite horizon, but below we focus on the infinite horizon problem under the average reward objective criterion and the discounted reward objective criterion.

2.1.1 Average Reward

To define such an objective we start a function for the value of a given state: $V_\mu : S \to R$. This represents the expected objective value obtained from S under μ. Further we define $V_\mu^{(t)}(s)$ to be the total expected reward after t transitions.

Exercise 23

Give an expression for the expected average reward after an infinite number of transitions.

End Exercise 23

For an *Ergodic* MDP, we therefore get:

$$V_\mu(s) = \sum_s R(s, \mu(s))\pi_\mu(s)$$

In other words, the sum of the expected rewards of all states. This expectation is calculated as the reward you get in a given state s under μ multiplied by the probability of being in that state.

Exercise 24

What is an Ergodic MDP?

End Exercise 24

2.1.2 Discounted Reward

We discount the value of a reward n steps away by a factor of γ^n for $0 < \gamma < 1$. Thus expected cumulative discounted reward is:

$$V_\mu(s) = E\left[R(s,\pi(s)) + \gamma R(s',\pi(s')) + \gamma^2 R(s'',\pi(s'')) + ...|\mu\right]$$

2.2 Solving MDPs

We specify a stochastic policy μ as matrix X whose entries $x(s,a)$ give the probability of taking action a in state s. Each row sums to 1, and each entry is in $[0,1]$. A deterministic policy enforces binary $x(s,a)$.

We solve in terms of $\pi(s,a)$, the steady-state probability of being in state s and taking action a.

Exercise 25

What is the difference between $x(s,a)$ and $\pi(s,a)$? How can you calculate $x(s,a)$ given $\pi(s,a)$?

End Exercise 25

2.2.1 Average Reward

With this background, we can formulate the following LP:

$$\begin{aligned}
\max \quad & \sum_s \sum_a R(s,a)\pi(s,a) && \text{Max Mean Expected Reward} \\
\text{s.t.} \quad & \sum_s \sum_a \pi(s,a) = 1 && \text{Unconditional prob. sum to 1} \\
& \sum_{a \text{ in } s'} \pi(s',a) = \sum_s \sum_{a \text{ going to } s'} \pi(s,a)P(s,a,s') \;\forall\, s' && \text{Prob Flow} \\
& \pi(s,a) \geq 0
\end{aligned}$$

Here we assume the Markov Chain for every policy is ergodic. Because of the set of constraints and irreducibility, the optimal $\pi(s,a)$ will turn out to induce binary values for all $x(s,a)$ – we can solve an LP, but will end up with a discrete solution.

2.2.2 Discounted Reward

First we choose a β vector such that $\sum_s \beta(s) = 1$ and $0 < \beta(s) \leq 1$. These represent the distribution of the start states (but needn't actually be them – they can be any value). Then we have:

$$\max \quad \sum_s \sum_a R(s,a)\pi(s,a) \qquad \text{Max Expected Discounted Reward}$$

$$\text{s.t.} \quad \sum_{a \text{ in } s'} \pi(s', a) - \gamma \sum_{s} \sum_{a \text{ going to } s'} \pi(s, a) P(s, a, s') = \beta(s') \; \forall \; s' \qquad \text{Discounted Reward Flow}$$

$$\pi(s, a) \geq 0$$

Here we **don't** have to assume the Markov Chain for every policy is ergodic, $\pi(s, a)$ will be independent of the choice of β. And still the $\pi(s, a)$ will turn out to induce binary values for all $x(s, a)$, despite the LP formulation.

Exercise 26

What do the $\pi(s, a)$ represent in the above formulation?

End Exercise 26

2.2.3 Other Techniques

There are actually a number of other ways to solve MDPs that are often used in practice. These include use of the Bellman Equations, and Iterative Solution Methods as hinted at in the lecture notes. These methods are outside the scope of this course, but ask offline if you want references.

2.3 Fundamental Theorem of MDPs

Fill in the blank

Exercise 27

Under the discounted reward criterion, the same policy μ^* is optimal for all ____________________.

End Exercise 27

2.4 Modeling

MDPs are widely used as models for dynamic processes. Lets try to model one:

Exercise 28

Swimmy the Fish has pretty good eye sight – at least for the important things in his little watery world: Food, Predators and Water (i.e. Nothing). His action space includes things like Pursuing Food, Swimming Away from predators, Waiting, etc. Model Swimmy as an MDP.

End Exercise 28

Exercise 29

Consider if Swimmy can also wait for the food to come to him when he sees it. How would this change the model?

End Exercise 29

Exercise 30

More generally suppose we were to add the action of Waiting regardless of the state. How would this change the model?

End Exercise 30

Exercise 31

What does a model like this one imply about feeding aquarium fish?

End Exercise 31

By completing the above exercises, you should now:

- understand the parts of a Markov decision processes and some different ways to evaluate the reward
- be able to model a dynamic processes with MDPs.

3 The Newspaper Business

Exercise 32

Raj, the owner of Out of Town News on Harvard Square, needs to decide how many Boston Globe newspapers he wants to buy in order to maximize profit. The problem is his business depends mostly on tourists, and the number of tourists depend on the weather of the day. The demand for newspaper, thus, is as follow:

	Nice weather	OK weather	Miserable weather
Demand	100	80	20
Probability	0.25	0.25	0.5

Each paper costs \$3 and can be sold for \$5. Any unsold paper is given away at the end of the day for free.

1. What is the expected demand?
2. If Raj has perfect information about the weather, how many newspaper should he buy?

End Exercise 32

By completing the above exercises, you should now:
- understand the basics of a two-stage optimization problem.

4 Optional: Two-stage optimization problem

4.1 Review

A 2-stage optimization problem is characterized by:

- A decision x, which has to be made in stage 1.
- An external uncertainty ξ, which is randomly distributed according to some known distribution, and can only be observed in stage 2. Note that, unlike in MDP, the realization of ξ is independent of x.
- A pay-off function (or cost function) $Q(x, \xi)$ that depends on the decision x that was undertaken and the realization of ξ. Note that sometimes obtaining $Q(x, \xi)$ requires taking an optimal recourse action $y(x, \xi)$ in stage 2.

Given a 2-stage optimization solution:

- The **omniscient solution** is the optimal solution when we have perfect information of how ξ will realize.
- The **expected value solution** is the optimal solution if we assume that ξ will be equal to the expected value of ξ.

4.2 Practice

Exercise 33

We continue with Raj's problem from exercise 1.

1. Which decision needs to be made in stage 1? Which external uncertainty happens in stage 2?
2. What is the recourse action in stage 2? What is Raj's pay-off?
3. Assume that Raj has perfect information about the weather, what is his expected profit?
4. Raj does not want to turn away his customer, so he decides to always buy 100 newspaper. What is his expected profit in this case?

5. It turns out that serving all customers is not a good business model. Raj turned to his son - a Yale student - for advice, and was told that he should just buy as many newspapers as the expected demand. What is his expected profit in this case?

End Exercise 33

By completing the above exercises, you should now:
- know how to find the omniscient solution and the expected value solution.

5 Optional: Optimal stochastic solution through stochastic programming

5.1 Review

Stochastic programming seeks to find the optimal stochastic solution through solving a 2-stage problem. Note that the lecture already shows the case of a minimization problem. In the case of a **maximization** problem, we have:

- Stage-one problem:

$$\max_x -c^T x + E_\xi[Q(x, \xi)]$$

  subject to

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}$$

  Here, $c^T x$ is the cost of taking action x in stage 1, and $E_\xi[Q(x, \xi)]$ is the expected pay-off in stage 2. for each possible value of ξ, $Q(x, \xi)$ is obtained by solving the stage 2 problem.

- Stage-two problem: For each value of ξ, we want to find

$$Q(x, \xi) = \max_y q^T y$$

  subject to

$$\begin{aligned} Tx + Wy &\leq h \\ y &\geq 0 \end{aligned}$$

  Here, recourse action y is subjected to some constraints that depend on the action x taken in the first stage, and we want to find the optimal $y(x, \xi)$ that maximizes the pay-off $Q(x, \xi) = \max q^T y$.

Please note that although we have different problems for each stage and for each value of ξ, in practice we **combine them into one LP** when formulating the problem. In addition to the farmer's problem in the lecture, we will also show you how to do this in the next exercise.

Using stochastic programming will give us the optimal stochastic solution. With that, we can calculate:

- The **Expected Value of Perfect Information** (EVPI) is the difference between the omniscient solution and the optimal stochastic solution. EVPI represents how much better we can do if we are to know exactly how ξ will realize in the second stage.

- The **Value of Stochastic Solution** is the difference between the optimal stochastic solution and the expected value solution. It shows how much stochastic programming improves our solution over naively using the expected value of ξ.

5.2 Practice

Exercise 34

Raj visited Yale during parents' weekend and sat in his son's optimization class. The experience led him to believe that it was a mistake asking for his son's advice. Once back in Cambridge, Raj found out that you are a student in AM121, and decided to ask for your help.

1. Using stochastic programming, what is Raj's stage 1 problem? Given a decision x, what is his stage 2 problem?
2. Formulate Raj's problem as a single LP.
3. Implement the problem in AMPL, and find the optimal solution.
4. Calculate EVPI and VSS?
5. An astrologist tells Raj that he can predict the weather, and offers to sell his prediction for \$100 a day. Should Raj accept the offer?

End Exercise 34

By completing the above exercises, you should now:
- understand the concepts of EVPI and VSS.
- know how to formulate and solve 2-stage stochastic optimization problems.

6 Solutions

Solution 1

A specification for the dynamics of a discrete system.

- A set of random variables S_t, one for each timestep. Each variable has a domain of the possible states $S = \{1, \ldots, m\}$ of the system:

$$S_t \in S \;\forall\; t \in \{0, 1, ...\}$$

- A transition probability for each variable S_t to each possible state S, given the realization of all previous states:

$$P(S_{t+1} = j | S_0 = a, S_1 = b, ..., S_t = i)$$

End Solution 1

Solution 2

The transition probabilities depend only on the state in the previous timestep:

$$P(S_{t+1} = j | S_0 = a, S_1 = b, ..., S_t = i) = P(S_{t+1} = j | S_t = i)$$

End Solution 2

Solution 3

The transition probabilities do not depend on time:

$$P(S_{t+1} = j | S_t = i) = p_{ij},$$

and independent of time period t.

End Solution 3

Solution 4

A **Markov Chain**

End Solution 4

Solution 5

Accessible

End Solution 5

Solution 6

Communicate, Accessible

End Solution 6

Solution 7

Irreducible

End Solution 7

Solution 8

Recurrent, Transient

End Solution 8

Solution 9

Absorbing

End Solution 9

Solution 10

Period

End Solution 10

Solution 11

Aperiodic

End Solution 11

Solution 12

Ergodic, Aperiodic. Additionally, note that all states of an irreducible Markov Chain are recurrent.

End Solution 12

Solution 13

No because you cannot get from 2 to 1.

End Solution 13

Solution 14

$\{\{1\}, \{2,3,4\}, \{5,6,7\}\}$

End Solution 14

Solution 15

$\{5,6,7\}$

End Solution 15

Solution 16

No because there is more than one class.

End Solution 16

Solution 17

No

End Solution 17

Solution 18

State 1 is aperiodic because there is no probability of reaching the state again.

State 5, 6, 7 are aperiodic because one can reach a state on consecutive steps (e.g. 5 and 6), and periodicity is a class property.

States 2, 3, and 4 are periodic with period 3 (even though the states are transient).

End Solution 18

Solution 19

No for several reasons. There are states in the chain that are:

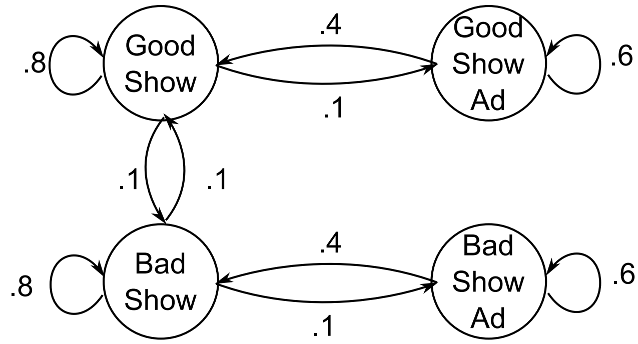
1. Periodic
2. Transient

Moreover, the Markov chain is not irreducible.

End Solution 19

Solution 20

We end up with a Markov Chain like the following:



End Solution 20

Solution 21

First we label the states:

- 1. Good show
- 2. Good show Ad
- 3. Bad show
- 4. Bad show Ad

We have the following transition matrix (remember p_{ij} denotes the probability of going from i to j):

$$\begin{bmatrix} .8 & .1 & .1 & 0 \\ .4 & .6 & 0 & 0 \\ .1 & 0 & .8 & .1 \\ 0 & 0 & .4 & .6 \end{bmatrix}$$

We have the following steady state equations $\pi_j = \sum_i \pi_i p_{ij}$:

$$\pi_2 = .1\pi_1 + .6\pi_2$$
$$\pi_2 = \frac{\pi_1}{4}$$

$$\pi_4 = .1\pi_3 + .6\pi_4$$
$$\pi_4 = \frac{\pi_3}{4}$$

$$\pi_1 = .8\pi_1 + .4\pi_2 + .1\pi_3$$
$$\pi_1 = .8\pi_1 + .1\pi_1 + .1\pi_3$$
$$\pi_1 = \pi_3$$

From the above we also have $\pi_2 = \pi_4$. Thus:

$$\begin{aligned}\pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \\ 2\pi_1 + 2\pi_2 &= 1 \\ \pi_1 &= \frac{1}{2} - \pi_2\end{aligned}$$

And substituting into the first equation, we get:

$$\begin{aligned}\pi_2 &= \frac{\pi_1}{4} \\ \pi_2 &= \frac{1}{4}(\frac{1}{2} - \pi_2) \\ \pi_2 &= \frac{1}{8} - \frac{\pi_2}{4} \\ \pi_2 &= \frac{4}{5}\frac{1}{8} = .1\end{aligned}$$

From which we get $(\pi_1, \pi_2, \pi_3, \pi_4) = (.4, .1, .4, .1)$. Notice that we could have also solved this by matrix inversion or any other linear algebra equation solving technique.

If you believe the initial probabilities, this implies there is a lot of junk on TV – and watching indiscriminately will expose you to it.

End Solution 21

Solution 22

In each state s, execute $\mu(s)$. Doing this in all states will reduce the MDP to a Markov Chain, with only the transitions stipulated by μ.

End Solution 22

Solution 23

$$V_\mu(s) = \lim_{t\to\infty} \frac{1}{t} V_\mu^{(t)}(s)$$

End Solution 23

Solution 24

An MDP for which every possible policy induces an ergodic Markov Chain (i.e. a chain that contains only one class and is aperiodic).

End Solution 24

Solution 25

$x(s, a)$ is conditioned on actually being in state s, whereas $\pi(s, a)$ is the unconditional probability. We thus have:

$$\pi(s) = \sum_a \pi(s, a)$$

and

$$\pi(s, a) = \pi(s)x(s, a)$$

Giving us:

$$x(s,a) = \frac{\pi(s,a)}{\pi(s)} = \frac{\pi(s,a)}{\sum_{a'} \pi(s,a')}$$

End Solution 25

Solution 26

The $\pi(s,a)$ represent the discounted expected amount of time being in state s and taking action a:

$$\pi(s,a) \equiv \sum_{t=0}^{\infty} \gamma^t \pi_t(s,a)$$

Where $\pi_t(s,a)$ is the probability of being in state s and taking action a at time t given initial state distribution β.

End Solution 26

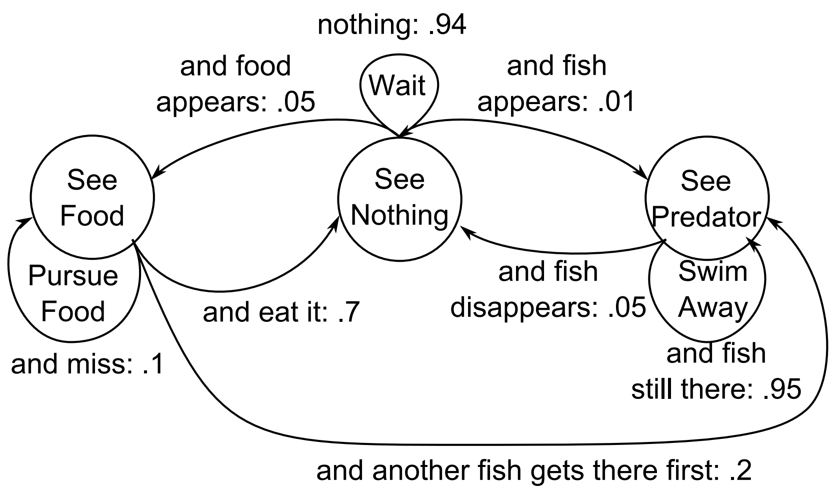
Solution 27

initial state distributions

End Solution 27

Solution 28

Suppose initially, Swimmy's life if boring and he follows the following Markov chain.



In reality we would imagine Swimmy to have other actions as well (e.g. other tactics when seeing a Predator, moving around instead of waiting, and so on).

End Solution 28

Solution 29

Swimmy will have this action available whenever he sees food. Waiting for Food would have the following possible results: Food takes a different route and Swimmy misses it (in which case he transitions into the see nothing state); Food continues along its current trajectory and Swimmy eats it; another fish can get to Food first.

End Solution 29

Solution 30

We have already described what would happen in the See Food state. In the See Predator state we may get the following results: fish (i.e predator) disappears; fish still there; fish eats Swimmy. The latter possibility introduces a new absorbing state: Dead!

End Solution 30

Solution 31

The optimal policy of this model is to always eat food if you see it. So if fish follow a MDP like this one, they will eat themselves to death (in which case you may want to model an absorbing dead state as well.) Don't over feed your fish!

End Solution 31

Solution 32

1. The expected demand is:

$$100 \times 0.25 + 80 \times 0.25 + 20 \times 0.5 = 55$$

2. Given perfect information about the weather, Raj should buy 100 papers on a day with nice weather, 80 newspaper on a day with OK weather, and 20 newspaper on a day with miserable weather.

End Solution 32

Solution 33

1. Raj's decision in stage 1 (x) is the number of newspapers to be bought. The external uncertainty (ξ) is the actual demand, which is only observed in stage 2.

2. We can consider the recourse action in this case to be how many newspapers to sell, given that we have x in our inventory and the actual demand is ξ. Because we are maximizing profit, the optimal action is simply to sell as many newspaper as possible, i.e. $\min(x, \xi)$. His pay-off is thus:

$$Q(x, \xi) = -3x + 5\min(x, \xi)$$

3. Assuming perfect information, Raj should buy 100 papers on a day with nice weather, 80 newspaper on a day with OK weather, and 20 newspaper on a day with miserable weather. Since he makes \$2 for each sold paper, his profit will be \$200, \$160, and \$40. The expected profit is thus

$$200\times\ 0.25\ +\ 160\times\ 0.25\ +\ 40\times\ 0.5 = \$110$$

   This is the **omniscient solution**.

4. If he buys 100 papers everyday, his profit will be:

   - On a nice day, the demand is 100, so he sells all 100 papers for a profit of

   $$100 \times \text{(5-3)} = \$200$$

   - On an OK day, the demand is 80, so he only sells 80 papers for a profit of

$$80 \times 5 - 100 \times 3 = \$ 100$$

- On a miserable day, the demand is only 20, so he only sells 20 papers for a profit of

$$20 \times 5 - 100 \times 3 = -\$200$$

His expected profit is thus:

$$200 \times 0.25 + 100 \times 0.25 + (-200) \times 0.5 = -\$25$$

5. As calculated in exercise 1, the expected demand is 55 papers. If he buys 55 papers everyday, his profit will be:

   - On a nice day, the demand is 100, so he sells all 55 papers for a profit of 55 × 2 = \$110.
   - On an OK day, the demand is 80, so he sells all 55 papers for a profit of 55 × 2 = \$110.
   - On a miserable day, the demand is only 20, so he only sells 20 papers for a profit of:

   $$20 \times 5 - 55 \times 3 = -\$65$$

   His expected profit is thus:

   $$110 \times 0.25 + 110 \times 0.25 + (-65) \times 0.5 = \$22.5$$

   This is the **expected value solution**.

End Solution 33

Solution 34

1. First, we introduce the following sets, parameters, and variables:

   Sets

S	set of possible weather

   Parameters

p	$i \in S$	probability of a particular weather type
d	$i \in S$	demand given a particular weather type
c		cost of buying a newspaper
r		revenue from selling a newspaper

   Variables

x	how many newspaper to buy (integer)
y_i for $i \in S$	how many newspaper is sold given a particular weather type (integer)
Q_i for $i \in S$	the revenue from selling y_i newspaper

   The **first-stage problem** is:

   Objective:

   $$\max_x -cx + \sum_{i \in S} p_i Q_i$$

   Constraints:

   $$x \geq 0$$

   Given a decision x in the first stage, the **second-stage problem** for each $i \in S$ is:

Objective:

$$Q_i = \max_{y_i} r y_i$$

Constraints:

$$\begin{aligned} y_i &\le x \\ y_i &\le d_i \\ y_i &\ge 0 \end{aligned}$$

2. The single LP problem that we need to solve is:

   Objective

$$\max -cx + \sum_{i \in S} p_i r y_i \tag{1}$$

   Constraints

$$\begin{aligned} y_i &\le d_i & \forall i \in S \quad (2) \\ y_i &\le x & \forall i \in S \quad (3) \\ x, y_i &\ge 0 & \forall i \in S \quad (4) \\ & & (5) \end{aligned}$$

3. The optimal solution is 20 newspaper for an expected profit of \$40.
4. EVPI is the difference between the omniscient solution and the optimal stochastic solution:

$$\text{EVPI} = 110 - 40 = \$70$$

   VSS is the difference between the optimal stochastic solution and the expected value solution:

$$\text{VSS} = 40 - 22.5 = \$17.5$$

5. Having a perfect weather prediction will improve Raj's expected daily profit by EVPI = \$70. Thus, Raj should not pay \$100 a day for the astrologist.

End Solution 34