Лекция 12

Марковские модели принятия решений

Марковские процессы

В данном разделе рассматриваются многостадийные задачи принятия решений с конечным числом состояний s_j ($j = 1, \ldots, m$) оптимизируемой системы S. Предполагается, что в дискретные моменты времени $t_1, t_2, \ldots$ система переходит в новое состояние в соответствии с некоторой матрицей переходных вероятностей:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}.$$

Элемент p_{ij} матрицы означает вероятность перехода системы из состояния s_i в состояние s_j. Таким образом, строки матрицы соответствуют "старым", а столбцы — "новым" состояниям. Очевидно, сумма элементов любой строки матрицы равна 1.

Такой процесс поведения системы называется *марковским*, если вероятность перехода системы в любое возможное состояние в каждый момент времени определяется только ее состоянием в предыдущий момент времени и не зависит от более ранней предыстории.

Марковские процессы

Пример Некоторая фирма занимается промышленной разработкой программного обеспечения для компьютерных систем. В начале каждого года она решает задачу замены оборудования, включающего технические и программные средства, используемые в производственном процессе и обеспечивающие необходимую технологическую среду разработки. В зависимости от результатов экспертной оценки оборудования состояние фирмы (это система S) оценивается как "хорошее" (1), "удовлетворительное" (2) и "плохое" (3). Следовательно, система может находиться в одном из трех указанных состояний. Матрица переходных вероятностей может иметь вид:

$$P^1 = \begin{bmatrix} 0,2 & 0,5 & 0,3 \\ 0 & 0,5 & 0,5 \\ 0 & 0 & 1 \end{bmatrix}.$$

Здесь, например, число 0, 3 означает, что если система находилась в "хорошем" состоянии, то в следующий момент изменения состояния (следующий момент анализа состояния фирмы) она окажется в "плохом" состоянии с данной вероятностью.

Марковские процессы

Предположим, что в зависимости от состояний, в которых последовательно оказывается система, может быть вычислен доход, приносимый фирмой. Логично предположить, что доход за период t_{i+1}—t_i зависит от уровня оснащенности фирмы современным оборудованием и технологическим окружением, включая профессиональный уровень персонала, также требующий непрерывного повышения.

Для моделирования этой ситуации можно матрице переходных вероятностей P^1 поставить в соответствие матрицу доходов R^1:

$$R^1 = \begin{bmatrix} 7 & 6 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

Элемент r_{ij} матрицы означает доход, полученный за период t_{i+1}—t_i при переходе системы из состояния i в состояние j. Так, число 5 означает доход, выраженный в некоторых условных единицах, при сохранении системой "удовлетворительного" состояния. Отрицательные значения отражают потери.

Марковские процессы

Имея матрицы P^1 и R^1, можно достаточно просто прогнозировать результаты функционирования системы. Ясно, что со временем оборудование устаревает и нуждается в обновлении в соответствии с новыми международными стандартами и требованиями рынка. В результате при постоянных матрицах P^1 и R^1 система может деградировать, неизменно оставаясь в плохом состоянии и принося одни убытки.

В реально работающих фирмах по результатам экспертного анализа проводится периодическое обновление оборудования с изменением технологического окружения и обучением персонала. Данный процесс моделируется изменением матриц переходных вероятностей и доходов.

Марковские процессы

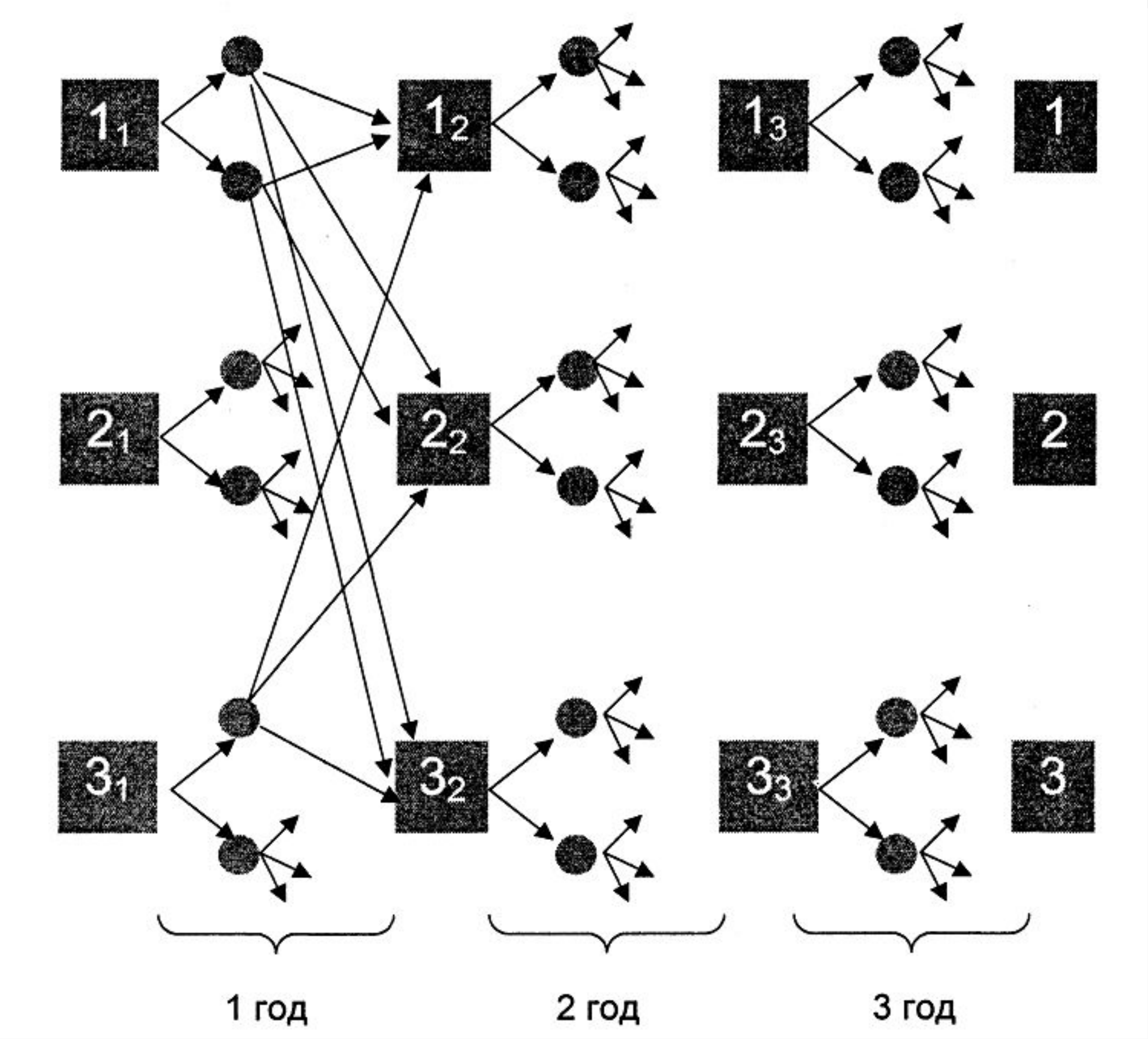
В нашем примере они, например, могут измениться следующим образом:

$$P^2 = \begin{bmatrix} 0{,}3 & 0{,}6 & 0{,}1 \\ 0{,}1 & 0{,}6 & 0{,}3 \\ 0{,}05 & 0{,}4 & 0{,}55 \end{bmatrix}; \qquad R^2 = \begin{bmatrix} 6 & 5 & -1 \\ 7 & 4 & 0 \\ 6 & 3 & -2 \end{bmatrix}.$$

Здесь в матрице доходов мы учли затраты на реорганизацию и модификацию. Например, элемент r_{11} матрицы R^2 оказывается меньше соответствующего элемента матрицы R^1.

На каждом этапе мы можем принять решение не проводить модернизацию фирмы и иметь матрицы P^1 и R^1 или принять решение о необходимых изменениях и получить матрицы P^2, R^2. Возникает проблема выбора или принятия решений с целью максимизации приносимого фирмой ожидаемого (речь идет о вероятностях!) дохода. Это многоэтапная задача принятия решений, т. к. выбор осуществляется каждый раз в заданные дискретные моменты времени.

Марковские процессы



Марковские процессы

Как обычно, квадратики означают решающие вершины. Каждый квадратик соответствует определенному состоянию системы в определенный момент времени. Знак i_j внутри квадрата означает, что в момент времени j, $j = 1, 2, 3$ (номер этапа) система находится в состоянии i, $i = 1, 2, 3$ (соответственно, "хорошее", "удовлетворительное" или "плохое" состояние). Две стрелки, исходящие из каждой "решающей" вершины, соответствуют двум альтернативам на каждом этапе: x_1 — проводить модернизацию (это верхняя стрелка, будем называть ее стрелкой или направлением 1) или x_2 — не проводить (это нижняя стрелка, будем называть ее стрелкой или направлением 2). Кружочки означают "случайные" вершины, переход из которых осуществляется в соответствии с выбранной матрицей переходных вероятностей.

Марковские процессы

Следуя общему алгоритму динамического программирования, решаем задачу с конца. Двигаемся справа налево по решающим вершинам. Начнем с вершины 1_3. Тогда при принятии решения x_1 (без модернизации) ожидаемый доход равен

$$d_1(1_3) = 0{,}2 \times 7 + 0{,}5 \times 6 + 0{,}3 \times 3 = 5{,}3.$$

При выборе x_2 (модернизация) имеем:

$$d_2(1_3) = 0{,}3 \times 6 + 0{,}6 \times 5 + 0{,}1 \times (-1) = 4{,}7.$$

Число 5,3 больше, чем 4,7, поэтому если мы окажемся в вершине 1_3, то пойдем по направлению 1, а сама вершина помечается числом 5,3. Стрелка 1 также выделяется.

Марковские процессы

Далее переходим к вершине 2_3. Получаем значения двух доходов в зависимости от принимаемых решений:

$$d_1(2_3) = 0{,}5 \times 5 + 0{,}5 \times 1 = 3{,}0;$$

$$d_2(2_3) = 0{,}1 \times 7 + 0{,}6 \times 4 + 0{,}3 \times 0 = 3{,}1.$$

Вершина 2_3 помечается числом 3,1 и выделяется направление 2.

Для вершины 3_3 получим:

$$d_1(3_3) = -1{,}0;$$

$$d_2(3_3) = 0{,}05 \times 6 + 0{,}4 \times 3 - 0{,}55 \times 2 = 0{,}4.$$

Вершина 3_3 помечается бóльшим числом 0,4 и выделяется стрелка 2.

Полученные числа 5,3, 3,1, 0,4 характеризуют один акт изменения состояния и получаемый при этом локальный доход. Далее эти вычисления уже не повторяются, а значения этих локальных доходов потребуются в дальнейших расчетах.

Марковские процессы

Переходим теперь к началу второго года. Начнем с вершины 1_2. При выборе направления (решения) 1 имеем:

$$d_1(1_2) = 0{,}2 \times 5{,}3 + 0{,}5 \times 3{,}1 + 0{,}3 \times 0{,}4 + 5{,}3 = 8{,}03.$$

Здесь число 5,3 отражает локальный доход этапа (рассчитанный ранее), а остальные слагаемые характеризуют наилучший ожидаемый доход, получаемый на оставшихся этапах. Для второго варианта решения для этой же вершины имеем:

$$d_2(1_2) = 0{,}3 \times 5{,}3 + 0{,}6 \times 3{,}1 + 0{,}1 \times 0{,}4 + 4{,}7 = 8{,}19.$$

Число 8,19 больше, чем 8,03, поэтому вершину 1_2 помечаем числом 8,19 и выделяем стрелку 2.

Марковские процессы

Для вершин 2_2 и 3_2 проводим аналогичные расчеты:

$$d_1(2_2) = 0 \times 5{,}3 + 0{,}5 \times 3{,}1 + 0{,}5 \times 0{,}4 + 3{,}0 = 4{,}75.$$

$$d_2(2_2) = 0{,}1 \times 5{,}3 + 0{,}6 \times 3{,}1 + 0{,}3 \times 0{,}4 + 3{,}1 = 5{,}61.$$

Выбираем число 5, 61 и выделяем стрелку 2. Далее имеем:

$$d_1(3_2) = 0 \times 5{,}3 + 0 \times 3{,}1 + 1 \times 0{,}4 + (-1) = -0{,}6.$$

$$d_2(3_2) = 0{,}05 \times 5{,}3 + 0{,}4 \times 3{,}1 + 0{,}55 \times 0{,}4 + 0{,}4 = 2{,}13.$$

Для первого этапа аналогично получаем:

$$d_1(1_1) = 10{,}38;$$

$$d_2(1_1) = 10{,}74;$$

$$d_1(2_1) = 6{,}87;$$

$$d_2(2_1) = 7{,}92;$$

$$d_1(3_1) = 1{,}13;$$

$$d_2(3_1) = 4{,}23.$$

Марковские процессы

Теперь обратная процедура динамического программирования закончена и, двигаясь от начала дерева решений к концу, можно "прочитать" оптимальное решение. А именно: числа 10,74, 7,92, 4,23 означают оптимальный ожидаемый доход, если, соответственно, система находилась первоначально в состояниях 1, 2 и 3. Эти ожидаемые доходы достигаются, если мы всегда будем вести себя "оптимально", т. е. в соответствии с помеченными на дереве решений стрелками. В частности, в каком бы состоянии мы ни находились в начале первого года, целесообразно решение, связанное с модернизацией оборудования. То же относится к началу второго года (все выделенные стрелки направлены "вниз"). И только если в начале третьего года мы окажемся в состоянии 1, нам нецелесообразно проводить модернизацию оборудования фирмы.