

НАУКОВІ ДОСЛІДЖЕННЯ

УДК 519.21

ЛИНЕАРИЗАЦИЯ МОДЕЛИ ПРОСТРАНСТВА СОСТОЯНИЙ СС

В. И. Большаков, д. т. н., проф., Ю. И. Дубров, д. т. н., проф.

«Время и пространство – это категории нашего мышления, а не условия нашего существования».

А. Эйнштейн

Ключевые слова: случайные факторы, функция отклика, экспертная оценка

Прежде всего, введем понятие линейного пространства. Идея линейности является одним из важнейших принципов математики. На этой основе построен классический анализ и вариационное исчисление. Более того, почти каждый физический процесс в малом (в определенном смысле) является линейным. Это позволяет нам делать достаточно точные выводы, изучая линейный, гораздо более простой для исследования объект. Если ракета пролетела за секунду 10 000 метров, то за две секунды она, вероятнее всего пройдет 20 000 метров. Если боковой ветер наклонил яхту на 40 градусов, то ветер, дующий вдвое слабее в том же направлении, наклонит яхту на 20 градусов. Отклик на малые воздействия линейно зависит от этих малых воздействий – таков естественнонаучный принцип, лежащий в основе огромного количества математических моделей. **Таким образом, линейное пространство – это понятие, обобщающее обычное (трехмерное) пространство в относительно малом диапазоне.**

Желая наиболее естественным образом распространить на общие линейные пространства возможность измерения, мы обратимся к существующей в обычном трехмерном пространстве операции скалярного произведения двух векторов. Скалярное произведение двух векторов есть произведение длин этих векторов и косинуса угла между ними. Откуда следует, что, зная скалярное произведение любой пары векторов, мы можем легко определить их длины и косинус угла между ними. Следовательно, в скалярном произведении потенциально заложены возможности измерения длин и углов. В общем, линейном пространстве нам легче будет сначала ввести понятие скалярного произведения, а затем с помощью этого понятия получить определения длины вектора и угла между векторами.

О п р е д е л е н и е 1. В линейном пространстве R определено скалярное произведение, если каждой паре векторов x, y поставлено в соответствие действительное число, обозначаемое (x, y) , причем это соответствие обладает следующими свойствами:

1. $(x, y) = (y, x)$ – симметрия;
2. $(\lambda x) y = \lambda(x, y)$, где λ – действительное число;
3. $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ – дистрибутивность;
4. для любого вектора, $(x, x) \geq 0$, $(x, x) = 0$ тогда и только тогда, когда $x = 0$.

О п р е д е л е н и е 2. Линейное пространство, в котором введено скалярное произведение, называется *евклидовым*.

Введение скалярного произведения – это, по сути, задание метрики, т. е. способа определения расстояния между двумя точками (элементами), или определение меры угла в той или иной геометрической системе.

Метрикой g на множестве U будем называть неотрицательную вещественную функцию, определенную на множестве всех упорядоченных пар элементов множества U и удовлетворяющую следующим условиям:

1. $g(x, y) = g(y, x)$;
2. $g(x, y) > 0$, если элементы x и y различны;
 $g(x, y) = 0$, если элементы x и y совпадают;
3. $g(x, y) \leq g(x, z) + g(z, y)$ (аксиома треугольника).

Аналогичным образом может быть установлено расстояние между любыми двумя

элементами любого множества¹.

Поскольку расстояние – это мера, определяемая между каждым двумя элементами любого множества произвольных элементов, постольку его формальное описание должно являться следствием физической природы этого множества. Выше мы уже отмечали, что часто математические модели, описывающие СС, удается аппроксимировать аналитически представляемыми геометрическими образами: n -мерными поверхностями, объемами. Поэтому всякая фигура в евклидовом пространстве может рассматриваться как метрическое пространство². Использование геометрической терминологии позволяет при исследовании весьма абстрактных объектов использовать громадный геометрический опыт, который приобретался нами при изучении пространства $E^n (n = 1, 2, 3)$. Такими объектами могут быть любые n -мерные поверхности и объемы, являющиеся геометрическими образами различных технологических процессов.

Пусть есть некоторый технологический процесс, математическая модель которого представлена некоторой поверхностью. Назовем расстоянием между двумя точками a и b , лежащими на этой поверхности, длину кратчайшей из дуг, соединяющей эти две точки и лежащей всеми своими точками на этой поверхности. Заданная таким образом метрика называется внутренней метрикой поверхности, а сама поверхность превращается в метрическое пространство. Откуда следует, что любая математическая модель включает внутреннюю метрику пространства, которую она аппроксимирует. Таким образом, внутренняя метрика пространства, – это закономерность, «заложенная» в математической модели, если эта модель адекватна исследуемому объекту или явлению.

Следует отметить, что большинство гипотез, встречающихся в прикладных и фундаментальных науках, могут интерпретироваться как некоторая модель. Однако, по мере увеличения сложности изучаемых объектов, стала возрастать сложность синтеза таких моделей, что привело к появлению задач, для решения которых существующие методы, например, математического программирования, оказываются непригодными. Последнее связано с тем, что их применение вызывает серьезные математические трудности, частично связанные с тем, что исследователи, создающие математическую модель интересующей их закономерности, исходят из предпосылки о том, что пространство состояний объекта идентификации однородное и изотропное³, т. е. имеет одинаковые свойства в каждой точке и по любому направлению. Это позволяет в дальнейшем вводить на всем пространстве единую метрику (как правило, евклидову).

Если однородность и изотропность пространства состояний СС подтверждаются экспериментами, то, как правило, этот факт можно объяснить технологическими ограничениями, диктующими проведение экспериментов в относительно малой (рабочей) области. Вероятно, поэтому эксперименты, рассматривающие те же СС в новых ситуациях, иногда даже незначительно отличающихся от предыдущих, часто оказываются в противоречии с выбранной моделью.

То обстоятельство, что в ряде задач удается ввести единую метрику на всем пространстве переменных, представляет собой исключение, так как выбор ее совершенно произволен. Тем не менее, некоторые из таких мер более полезны и интуитивно более оправданы, чем другие, так как «...некоторые силы в природе следуют одной, другие своей особой геометрией».

Вероятно, для того, чтобы модель СС адекватно отражала ее свойства, необходимо отказаться от предположения об изотропности пространства состояний и для каждой области этого пространства нужно создавать свою математическую модель. Для этого в каждой области пространства состояний, заданной ограничениями переменных, должно быть введено свое «меропределение», или «установлена своя метрика».

Если речь идет о числах, например x и y , то под расстоянием между ними естественно понимать абсолютную величину их разности, что в геометрическом смысле эквивалентно расстоянию между соответствующими точками L_1 и L_2 на числовой прямой:

$$L_1 L_2 = |x - y|.$$

¹ Подробней по данным вопросам можно обратиться к специальной литературе (см. например [22]).

² Метрическим пространством называется совокупность точек, в которой для каждой двух точек определено положительное расстояние.

³ Изотропность [от гр. isos – равный, одинаковый, подобный + tropos свойство]. Как правило, характеризует одинаковость свойств (пространства, вещества и др.).

Если речь идет о парах чисел (x_1, y_1) и (x_2, y_2) , то, представляя их в виде точек, в декартовой системе координат расстояние задается формулой:

$$L_1 L_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (1)$$

Расстояние между точками L_1 и L_2 в пространстве n -измерений определяется как :

$$L_1 L_2 = \sqrt{(x_1 - y_1)^2 + (x_2 + y_2)^2 + \dots + (x_n - y_n)^2}. \quad (2)$$

Отметим, что метрика, определяемая приведенными формулами, обусловлена используемыми геометрическими представлениями и не является логически необходимой. Она носит название евклидовой метрики.

Поясним это на следующем примере. Представим себе произвольную кривую поверхность в пространстве с нанесенной на ней произвольной криволинейной координатной сеткой.

Расстояние между точками L_1 и L_2 , на любой поверхности обыкновенно определяется как длина кратчайшей из всех линий, которые могут быть проведены на поверхности от одной точки к другой. Это расстояние есть более или менее сложная функция от координат этих точек. Оно вовсе не обязательно совпадает с расстоянием между «изображениями» соответствующих точек на плоскости.

Поэтому при построении математической модели СС надо четко разделять два вопроса: какие свойства абстрактного пространства следуют из некоторой системы вводимых нами аксиом и какими свойствами обладает реальный объект, идентификацию которого мы производим посредством моделирования его пространства состояний. Задавая систему аксиом, мы собственно задаем свойства пространства состояний и, задавая ее однажды, мы, иногда ошибочно, порываем эту связь, поскольку дальше абстрактное математическое пространство начинает «жить» самостоятельной жизнью, развиваясь по законам логики, которые за пределами некоторых технологических ограничений могут быть неадекватными реальным процессам, объектам или явлениям.

Подытожим сказанное. Всякое приложение математики к естественным наукам включает построение абстрактной математической модели, описывающей интересующий нас объект или явление. Данная математическая модель, заданная, например, в виде функции, включает в неявном виде внутреннюю метрику пространства состояний объекта идентификации, поскольку согласно этой модели всегда можно вычислить расстояние между точками (численными значениями функции, координаты которой заданы значениями независимых переменных). Затем, развивая эту модель, мы получаем следствия, сравнивая которые с экспериментом, улучшаем модель (например, уточняем коэффициенты при неизвестных). Этим самым мы уточняем метрику пространства состояний, получаем новые следствия, опять сравниваем их с реальным процессом и т. д. Конечно, это всего лишь грубая схема, которая нам нужна лишь только для того, чтобы понять «в первом приближении» связь между процессом построения и улучшения математической модели реального явления и задания метрики пространства переменных и ее уточнения.

Еще одно осложнение, возникающее при идентификации метрики пространства состояний СС, заключается в том, что на одном и том же множестве U , содержащем более одной точки, может быть задано бесконечно большое число различных метрик, из которых надо выбрать одну, наилучшим образом отвечающую моделируемому явлению. Но допустим, что такой выбор сделан. Тогда может возникнуть еще одна особенность выбранного пространства состояний. Оно, это пространство, может деформироваться и дрейфовать под действием неизвестных, и поэтому не включенных в модель факторов, спонтанное и неконтролируемое изменение которых приводит к изменению метрики во всей или в части его рабочей области. Представляется, что такие изменения можно моделировать переходом от представления математической модели некоторой функцией f , к ее представлению некоторой функцией w , что, в свою очередь, позволяет поставить вопрос о выяснении расстояния между данными функциями, т. е. выяснении вновь полученной структуры пространства состояний с мерой на нем.

Возвратимся к интерпретации пространства U состояний объекта идентификации некоторой кривой поверхностью с нанесенной на нее произвольной криволинейной координатной сеткой. Рассматривая одну криволинейную линию этой сетки, например, как функцию $f(x)$ одной независимой переменной x , заданной в промежутке $I(a \leq x \leq b)$, а другую криволинейную линию этой же сетки – как функцию $w(x)$, заданную в этом же

промежутке I , введем расстояние между функциями $f(x)$ и $w(x)$. При этом для простоты принимаем, что речь идет только о непрерывных функциях.

В качестве «меры отклонения» функции $f(x)$ от функции $w(x)$ в промежутке I можно принять максимум абсолютной величины их разности в этом промежутке:

$$g(f, w) = \max\{|f(x) - w(x)|\}.$$

Исходя из знаний предметной области описываемой функциями f и w , можно было бы в качестве метрики $g(f, w)$ взять один из интегралов

$$\int_a^b |f(x) - w(x)| dx \quad \text{или} \quad \int_a^b [f(x) - w(x)]^2 dx.$$

Как правило, нарушение условия адекватности функции $f(x)$ объекту идентификации приводит к тому, что, подбирая конечное число экспериментальных точек (например, численных значений этой функции и соответствующих им значений независимых переменных – x), подбирают так элементарную непрерывную функцию (чаще всего рациональный многочлен $P(x)$), чтобы график этой функции проходил точно через выбранные точки.

Но, как это отмечалось выше, трудность идентификации сложных объектов заключается в том, что графиков функций, проходящих точно через выбранные точки, можно провести бесчисленное множество, так же, как можно предполагать бесчисленное множество метрик, формально адекватно описывающих объект. В этой связи можно предложить, чтобы выбор той или иной метрики пространства состояний объекта идентификации производился с учетом эвристических процедур, в которых, так или иначе, учитываются априорные знания исследователя об объекте идентификации.

Еще раз отметим, что закономерности, сформированные в виде математических моделей, нужны главным образом для того, чтобы делать предсказания будущих событий. Эти предсказания, вероятно, могут быть абсолютно точны только при условии изоморфизма⁴ метрик математической модели и объекта.

Синтез моделей, удовлетворяющих этому требованию, теоретически возможен с применением тензорного анализа. Однако для неэлементарного многообразия⁵, которым можно интерпретировать область переменных SS , невозможно в целом ввести координатную систему с обычными требованиями взаимной однозначности и непрерывности соответствия. Поэтому вся геометрия многообразия должна быть извлечена из задания в нем множества координатных систем, связанных между собой произвольными взаимно однозначными преобразованиями, что для практических применений представляется чрезвычайно сложным.

Экспертные системы (ЭС) – это яркое и быстро прогрессирующее направление в области искусственного интеллекта (ИИ). Причиной повышенного интереса, который ЭС вызывают к себе на протяжении всего времени своего существования, является возможность их применения к решению задач из самых различных областей человеческой деятельности. Пожалуй, не найдется такой проблемной области, в которой не было бы создано ни одной ЭС или, по крайней мере, такие попытки не предпринимались бы.

ЭС – это набор программ или программное обеспечение, которое выполняет функции эксперта при решении какой-либо задачи в области его компетенции. ЭС, как и эксперт-человек, в процессе своей работы оперирует со знаниями. Знания о предметной области, необходимые для работы ЭС, определенным образом формализованы и представлены в памяти ЭВМ в виде базы знаний, которая может изменяться и дополняться в процессе развития системы [27].

ЭС выдают советы, проводят анализ, выполняют классификацию, дают консультации и ставят диагноз. Они ориентированы на решение задач, обычно требующих проведения экспертизы человеком-специалистом. В отличие от машинных программ, использующих процедурный анализ, ЭС решают задачи в узкой предметной области (конкретной области

⁴Изоморфизм – важное понятие современной математики, подробное определение которому можно найти в специальной литературе, а для рассматриваемого случая отметим, что два изоморфных множества неразличимы по своим свойствам в терминах понятий, их определяющих.

⁵Многообразие – математическое понятие, обобщающее на n -мерный случай понятие поверхности без пересечений. Риманово многообразие или риманова поверхность аналитической функции $w = f(z)$ комплексного переменного z – поверхность R такая, что данная полная аналитическая функция, вообще говоря, многозначная, может рассматриваться как однозначная аналитическая функция $w = F(p)$ точки p поверхности R .

экспертизы) на основе дедуктивных рассуждений. Такие системы часто оказываются способными найти решение задач, которые не структурированы и плохо определены. Они справляются с отсутствием структурированности путем привлечения эвристики, т. е. правил, взятых «с потолка», что может быть полезным в тех системах, когда недостаток необходимых знаний или времени исключает возможность проведения полного анализа.

Главное достоинство ЭС – возможность накапливать знания, сохранять их длительное время, обновлять и тем самым обеспечивать относительную независимость конкретной организации от наличия в ней квалифицированных специалистов. Накопление знаний позволяет повышать квалификацию специалистов, работающих на предприятии, используя наилучшие, проверенные решения.

Практическое применение искусственного интеллекта на машиностроительных предприятиях и в экономике основано на ЭС, позволяющих повысить качество и сохранить время принятия решений, а также способствующих росту эффективности работы и повышению квалификации специалистов.

Отличие ЭС от других программных продуктов. Основными отличиями ЭС от других программных продуктов являются использование не только данных, но и знаний, а также специального механизма вывода решений и новых знаний на основе имеющихся. Знания в ЭС представляются в такой форме, которая может быть легко обработана на ЭВМ. В ЭС известен алгоритм обработки знаний, а не алгоритм решения задачи. Поэтому применение алгоритма обработки знаний может привести к получению такого результата при решении конкретной задачи, который не был предусмотрен. Более того, алгоритм обработки знаний заранее не известен и строится по ходу решения задачи на основании эвристических правил. Решение задачи в ЭС сопровождается понятными пользователю объяснениями. Качество получаемых решений обычно не хуже, а иногда и лучше достигаемого специалистами. В системах, основанных на знаниях, по которым решаются проблемы в конкретной предметной области, хранятся в базе знаний. Проблемы ставятся перед системой в виде совокупности фактов, описывающих некоторую ситуацию, и система с помощью базы знаний пытается вывести заключение из этих фактов. Качество ЭС определяется размером и качеством базы знаний (БЗ). В любой момент времени в системе существуют три типа знаний:

структурированные знания – статические знания о предметной области. После того как эти знания выявлены, они уже не изменяются;

структурированные динамические знания – изменяемые знания о предметной области. Они обновляются по мере выявления новой информации.

рабочие знания – знания, применяемые для решения конкретной задачи или проведения консультации.

Все перечисленные выше знания хранятся в базе знаний. Для ее построения требуется провести опрос специалистов, являющихся экспертами в конкретной предметной области, а затем систематизировать, организовать и снабдить эти знания указателями, чтобы впоследствии их можно было легко извлечь из базы знаний.

Структура традиционной экспертной системы. Традиционная ЭС состоит из следующих компонентов:

База знаний предназначена для хранения экспертных знаний о предметной области, используемых при решении задач экспертной системой. База знаний состоит из набора фреймов и правил-продукций.

Фреймы используются в базе знаний для описания объектов, событий, ситуаций, прочих понятий и взаимосвязей между ними. Фрейм – это структура данных, состоящая из слотов (полей) [31].

Правила используются в базе знаний для описания отношений между объектами, событиями, ситуациями и прочими понятиями. На основе отношений, задаваемых в правилах, выполняется логический вывод. В условиях и заключениях правил присутствуют ссылки на фреймы и их слоты.

База данных предназначена для временного хранения фактов или гипотез, являющихся промежуточными решениями или результатом общения системы с внешней средой, в качестве которой обычно выступает человек, ведущий диалог с экспертной системой.

Машина логического вывода – механизм рассуждений, оперирующий знаниями и данными с целью получения новых данных из знаний и других данных, имеющихся в рабочей памяти. Для этого обычно используется программно реализованный механизм дедуктивного логического вывода (какая-либо его разновидность) или механизм поиска решения в сети

фреймов или семантической сети.

Машина логического вывода может реализовывать рассуждения в виде:

- дедуктивного вывода (прямого, обратного, смешанного);
- нечеткого вывода;
- вероятностного вывода;
- унификации;
- поиска решения с разбиением на последовательность подзадач;
- поиска решения с использованием стратегии разбиения пространства поиска с учетом уровней абстрагирования решения или понятий, с ними связанных;
- монотонного или немонотонного рассуждения,
- рассуждений с использованием механизма аргументации;
- ассоциативного поиска с использованием нейронных сетей;
- вывода с использованием механизма лингвистической переменной.

Подсистема общения служит для ведения диалога с пользователем, в ходе которого ЭС запрашивает у пользователя необходимые факты для процесса рассуждения, а также дающая возможность пользователю в какой-то степени контролировать и корректировать ход рассуждений экспертной системы.

Подсистема объяснений необходима для того, чтобы дать возможность пользователю контролировать ход рассуждений и, может быть, учиться у экспертной системы. Если нет этой подсистемы, экспертная система выглядит для пользователя как «вещь в себе», решениям которой можно либо верить, либо нет. Нормальный пользователь выбирает последнее, и такая ЭС не имеет перспектив для использования.

Подсистема приобретения знаний служит для корректировки и пополнения базы знаний. В простейшем случае это – интеллектуальный редактор базы знаний, в более сложных экспертных системах – средства для извлечения знаний из баз данных, неструктурированного текста, графической информации и т. д.

Здесь необходимо кратко напомнить о нетрадиционной ЭС.

База знаний (БЗ). База знаний (БЗ) экспертных систем предназначена для хранения данных. Данные в рабочей памяти могут быть однородны или разделяются на уровни по типам данных.

В последнем случае на каждом уровне БЗ хранятся данные соответствующего типа. Выделение уровней усложняет структуру экспертной системы, но делает систему более эффективной [27].

В современных экспертных системах данные в БЗ рассматриваются как изолированные или как связанные. В первом случае рабочая память состоит из множества простых элементов, а во втором – из одного или нескольких (при нескольких уровнях в БЗ) сложных элементов (например, объектов). При этом сложный элемент соответствует множеству простых, объединенных в единую сущность. Теоретически оба подхода обеспечивают полноту, но использование изолированных элементов в сложных предметных областях приводит к потере эффективности.

Данные БЗ в простейшем случае являются константами и (или) переменными. При этом переменные могут трактоваться как характеристики некоторого объекта, а константы – как значения соответствующих характеристик. Если в БЗ требуется анализировать одновременно несколько различных объектов, описывающих текущую проблемную ситуацию, то необходимо указывать, к каким объектам относятся рассматриваемые характеристики. Одним из способов решения этой задачи является явное указание того, к какому объекту относится характеристика [28].

Если БЗ состоит из сложных элементов, то связь между отдельными объектами указывается явно. При этом каждый объект может иметь свою внутреннюю структуру. Необходимо отметить, что для ускорения поиска и сопоставления данные БЗ могут быть связаны не только логически, но и ассоциативно.

Отличительные особенности разных экспертных систем. 1) Традиционная ЭС – это система, созданная с помощью формализации, т. е. системы «вопрос – ответ». Данные экспертные системы должны иметь (имеют) очень обширную базу знаний, чтоб учесть (учитывать) все возможные ситуации. По опыту создания данных программ необходимо от 3 до 10 лет и создание их должно производиться при участии большого количества специалистов в данной конкретной области.

2) Однопараметрическая ЭС – это система, у которой базой знаний являются зависимости

