ЗАГАДКИ ДИФФЕРЕНЦИРОВАНИЯ

С.Я. Серовайский

Казахский национальный университет им. аль-Фараби, Казахстан, Алматы

Предметом нашего обсуждения является замечательная операция дифференцирования, являющаяся одним из центральных понятий математического анализа. Глубокие идеи дифференцирования проходят практически через все направления математики и находят своё применение за ее пределами. Мы рассмотрим наиболее яркие моменты эволюции взглядов на производную от ее ранних приложений в геометрии и механики до конструкций обобщенной производной и дифференцирования операторов, относящихся к двадцатому веку и во многом определивших развитие современной математики.

Три лика производной

В окружающем нас мире всё неуклонно меняется. И для того чтобы хоть как-то понять динамику протекающих событий, хотелось как-то оценить скорость происходящих изменений. Отталкиваясь от законов природы, мы можем сопоставить исследуемому процессу его *математическую модель*. Эта модель призвана описать характер изменения состояния рассматриваемой системы в силу тех или иных причин, а вот скорость изменения каких-либо величин описывается с помощью одного из важнейших математических понятий – производной.

Производная имеет три разных интерпретации. С позиций *анализа* предметом исследования является *функция* $x = x(t)$. Пусть для двух аргументов t и $t+\tau$ она принимает некоторые значения $x(t)$ и $x(t+\tau)$. Теперь можно вычислить отношение приращения функции $x(t+\tau)-x(t)$ к приращению аргумента τ и попытаться перейти к пределу, когда величина τ стремится к нулю. Если таковой предел существует, то он и называется *производной* функции x в точке t и обозначается через $x'(t)$. Таким образом, аналитически производная характеризуется *пределом*

$$x'(t) = \lim_{\tau\to 0}\frac{x(t+\tau)-x(t)}{\tau}.$$

В *геометрии* рассматриваемым объектом оказывается *кривая*. В плоскости переменных t и x определяются две точки на кривой – A с координатами t и $x(t)$, а также B с координатами $t+\tau$ и $x(t+\tau)$ (рис. 1). В соответствующем прямоугольном треугольнике ABC катеты BC и AC имеют длины $x(t+\tau)-x(t)$ и τ. Их отношение есть тангенс угла α наклона секущей AB рассматриваемой кривой. Теперь при стремлении τ к нулю, когда точка B стремится к A, секущая постепенно преобразуется в касательную. Тем самым с позиций геометрии производная определяется как тангенс угла наклона *касательной* к кривой в данной точке.

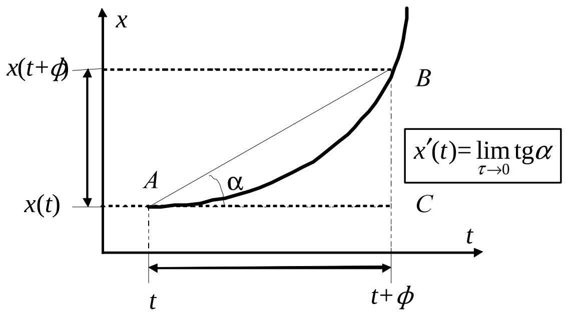

Рис. 1. Геометрический смысл производной.

В *механике* же рассматривается прямолинейное *движение* некоторого тела. В момент времени t тело находилось в точке $x(t)$, а в момент времени $t+\tau$ оно переместилось в точку $x(t+\tau)$. Таким образом, за время τ тело прошло путь длиной $x(t+\tau)-x(t)$. Средняя скорость движущегося тела на заданном интервале времени представляет собой отношение пройденного пути ко времени движения. При стремлении длины временного интервала τ к нулю мы получаем скорость движущегося тела в момент времени t. Тем самым в рамках механики производная оказывается мгновенной *скоростью*.

Многоликость понятия производной предъявляла повышенные требования к исследователям-первопроходцам. Это должны были быть специалисты, в совершенстве владеющие разными предметными областями. Поэтому кажется вполне естественным, что у истоков понятия производной стоял именно **Исаак Ньютон** (рис. 2), бывший в равной степени великим физиком и математиком. Для основоположника исчисления бесконечно малых и классической механики анализ, геометрия и механика были ведущими направлениями исследований. Физик Ньютон изучал движение тел и рассчитывал соответствующие скорости. Аналитик Ньютон разработал *метод флюксий*, под которыми понимаются скорости изменения переменных величин, т.е. те же производные. Ну а геометрия была для него надежным средством наглядной иллюстрации разрабатываемых механических и аналитических конструкций.

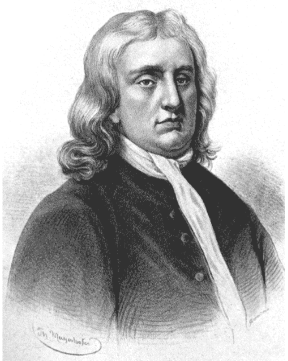

Рис. 2. Исаак Ньютон (1643 – 1727)

Принято считать, что производная была определена Ньютоном во второй половине 17 века. Однако это понятие имеет свою предысторию. Не зря сам Ньютон говорил: *"Я видел далеко, поскольку стоял на плечах гигантов"*. Одним из таких гигантов был, **Рене Декарт** (рис. 3), выдающийся французский мыслитель первой половины 17 века.

Декарт, основоположник философии нового времени, был создателем философской теории естествознания. Согласно Декарту весь мир представляет собой хорошо отлаженный механизм, а любая наука сводится

Рис. 3. Рене Декарт (1596 – 1650)

к механике. Языком науки является математика, призванная описывать любые явления окружающего мира, но поскольку явления эти происходят в пространстве, то непременно должно существовать надежное средство для описания пространственных форм.

Декарт вводит понятие *переменной величины*, которой можно дать различную интерпретацию. Например, это может быть траектория движущегося тела, что относится к механике (кстати, к закону инерции или первому закону Ньютона Декарт пришел еще до самого Ньютона). В анализе, переменной величине естественно соответствует понятие *функции*. Мы еще знаем Декарта и как основоположника *аналитической геометрии*, создателя *метода координат*. Идея этого метода состоит в том, что любой геометрический объект может быть описан аналитически, поскольку каждая его точка однозначно характеризуется своими координатами. Понятно, что за тремя интерпретациями переменной величины у Декарта нетрудно увидеть три интерпретации производной у Ньютона.

Но и Декарт творил не на пустом месте. Так, его аналитическим разработкам предшествовали: функциональные зависимости Франсуа Виета (конец 16 века); элементы классификации функций Николя Орема (14 век); таблицы тригонометрических функций Клавдия Птолемея (второй век); простейшие таблицы квадратов и кубов (древний Вавилон второе тысячелетие до нашей эры).

Построением касательных занимались когда-то греческие математики. Так Евдокс

(IV век до нашей эры) определяет касательную к кругу, его последователь Архимед – к параболе, гиперболе и спирали, а Аполлоний описывает касание окружностей. Много позднее Николя Орем строит графики функций. Иоганн Кеплер, Жиль Роберваль и Эванжелиста Торричелли (XVII век) уже владеют достаточно совершенной техникой построения касательных к различным объектам, а чуть позднее Пьер Ферма и Блез Паскаль одновременно с Декартом при построении касательных используют аналитический аппарат.

Элементы механики можно обнаружить уже в знаменитых апориях Зенона, великий философ Аристотель (IV век до нашей эры) дает количественное определение скорости. С понятием мгновенной скорости работает арабский мыслитель Сабит ибн Корра (IX век). Англичанин Томас Брадвардин (XIV век) устанавливает связь скорости с силой, француз Жан Буридан определяет понятие импульса, а Николя Орем исследует мгновенную скорость и равноускоренное движение. Ну и, конечно же, прямыми предшественниками Ньютона на пути построения классической механики и разработки ее математического аппарата были Кеплер и Галилей.

Как мы видим, немалые шансы определить производную имел, возможно **Николя Орем**, отчетливо понимавший глубокую связь между механикой, анализом и геометрией. Однако для адекватного описания рассматриваемых явлений ему определенно не хватало математической символики, появившейся лишь через двести лет у Франсуа Виета, и в еще большей степени – у Декарта, и уже практически в современном виде – у Лейбница. **Иоганн Кеплер** и **Галилео Галилей** подобным аппаратом уже владели неплохо, и подошли к открытию гораздо ближе, но они еще не чувствовали глубокой связи между анализом и геометрией. Подобную связь открыли математики последующего поколения –Декарт, Пьер Ферма. Казалось бы, этим выдающимся мыслителям ничего не мешало сделать последний решающий шаг, и шаг-то казался очень малым... Но вот сделать его бесконечно малым, т.е. попросту перейти к пределу в соответствующем соотношении они так и не догадались. Вот тогда и пришел Ньютон… Именно его и считают основоположником классического математического анализа... Но не его одного…

Дифференциальное исчисление

В истории математики Ньютон оказался неразрывно связанным со своим современником – великим немецким математиком и столь же великим философом **Готфридом Вильгельмом Лейбницем** (рис. 4). Но только ли математиком и философом? Он был филологом и геологом, инженером и дипломатом, конструктором одного из первых вычислительных устройств и биологом-эволюционистом. Трудно найти такую область знания, к которой бы не прикоснулась умелая рука этого блистательного мастера.

Но главное, конечно, математика…, практически вся терминология и символика классического математического анализа восходят к Лейбницу. Он занял прочное место в истории математики не только за счет совершенной формы, но и благодаря не менее яркому содержанию. К своим изысканиям Лейбниц приступил явно позднее Ньютона, однако, опубликовал свои работы раньше него, и шел он к цели несколько иным путем. Если для Ньютона важнейшим было понятие производной, то в анализе Лейбница центральную роль – дифференциал.

Подобно Ньютону Лейбниц рассматривает приращение $x(t+\tau)-x(t)$ функции $x=x(t)$ в точке t, соответствующее отрезку BC (рис. 5). При определенных условиях это приращение можно представить в виде суммы $a\tau+b(\tau)$, где a есть некоторое число, а функция $b=b(\tau)$ такова, что отношение $b(\tau)/\tau$ стремится к нулю при τ стремящемся к нулю. Тогда функция x называется *дифференци-*

руемой в точке t, а величина $a\tau$ – *дифференциалом* рассматриваемой функции в ука-

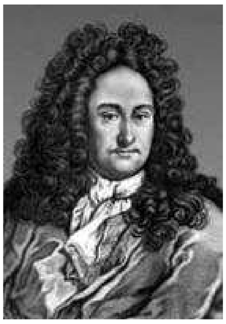

Рис. 4. Готфрид Вильгельм Лейбниц (1646 – 1716)

занной точке. Очевидно, число a есть не что иное, как производная $x'(t)$ данной функции, а описанное приращение характеризуется равенством

$$x(t+\tau)-x(t)=x'(t)\tau+b(\tau).$$

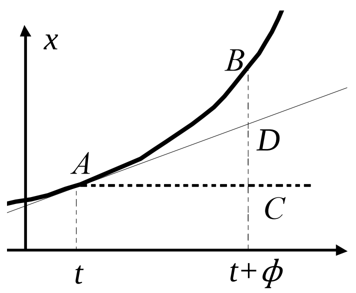

Рис. 5. Понятие дифференциала.

Дифференциал функции (ему соответствует отрезок *CD* на рисунке) оказывается тем самым главной линейной частью приращения. Линейной, поскольку таковой оказывается зависимость дифференциала от приращения аргумента τ (этой зависимости соответствует прямая, проходящая через точки *A* и *D*), а главной, потому что второе слагаемое в правой части последнего равенства (ему соответствует отрезок *BD* на рисунке) при малом τ оказывается, как говорят математики, *величиной более высокого порядка малости* (в этом и заключается смысл ограничения, налагаемого на функцию b). С подачи Лейбница процедура вычисления производной функции была названа *дифференцированием*, а соответствующий раздел *математического анализа* – *дифференциальным исчислением*.

Справедливости ради следует отметить, что понятием дифференциала в определенной степени владел еще **Блез Паскаль**, современник Ферма и Декарта. Однако подобно им и в отличие от Лейбница Паскаль не допускает приращение аргумента бесконечно малой величиной. Ну а современное определение дифференциала дает через сто лет после Лейбница Жозеф Луи Лагранж.

Следует отметить, что теорию Ньютона и Лейбница вообще-то нельзя считать строгой, т.к. у них отсутствовало понятие *предела*, а без этого ни производную, ни дифференциал вообще-то определить невозможно. Интуитивно они чувствовали нечто подобное, допуская приращение аргумента сколь угодно малым, но охарактеризовать предел аналитически не смогли. Таковое определение дали в начале XIX века **Бернард Больцано** и **Луи Огюстен Коши**. Они определили производную на основе теории пределов в форме, практически не отличающейся от современной. Именно им и выпала честь завершить построение фундамента дифференциального исчисления, а заодно всего классического математического анализа. Однако в основаниях теории предела оставался еще один серьезный пробел, на который обратил внимание **Карл Вейерштрасс** (конец XIX века).

Понятие предела вводится для действительных чисел. Областью определения и областью значений дифференцируемой функции также выступают множества действительных чисел. Однако строгое определение *действительного числа* в математике на тот момент отсутствовало, и лишь после того, как Вейерштрасс и его сподвижники охарактеризовали действительные числа безупречно строго, теория предела и весь классический математический анализ, включая теорию дифференцирования, приняли более или менее законченный вид.

Далее вставал вопрос, насколько широк класс функций, обладающих производными? Поскольку первооткрыватели и их ближайшие последователи работали с достаточно хорошими функциями, поначалу казалось, что всякая непрерывная функция должна непременно оказаться дифференцируемой. Однако вскоре выяснилось, что существуют и недифференцируемые непрерывные функции, например, модуль $x(t) = |t|$, равный t для положительных значений аргумента и $-t$ для его отрицательных значений (к рассмотрению этого удивительного объекта мы еще вернемся). Позднее Больцано, а вслед за ним и Вейерштрасс, привели примеры всюду непрерывных функций, которые не имеют производных ни в одной точке.

Спрашивается, а зачем вообще нужны столь экзотические объекты? Но оказалось, что непрерывные функции, как правило, таковыми и являются. Другими словами, непрерывные функции, обладающие производными хоть где-нибудь, это исключение, а не правило. В начале 20 века непрерывные нигде не дифференцируемые функции нашли важное практическое приложение. Так, в броуновском движении каждая частица постоянно меняет направление своего движения в результате соударения с другими частицами. Ее движение оказывается случайным (более точно, характеризуется *случайным процессом Винера*). Понятно, что траектория броуновской частицы нигде не прерывается, т.е. мы имеем дело с непрерывной функцией времени. Однако, учитывая, что время движения частицы между двумя последовательными соударениями чрезвычайно мало, то можно считать, что в каждый момент времени она сталкивается с другой частицей. Направление движения рассматриваемой частицы за счет этого меняется, а значит, происходит скачкообразное изменение вектора скорости. Тем самым, функция, характеризующая траекторию движения частицы, в данный (а значит, и в любой другой) момент времени производную не имеет.

Надо отметить, что становление дифференциального исчисления изначально было обусловлено ее высокой практической значимостью, и дальнейшее развитие этой теории неразрывно связано с потребностями ее многочисленных приложений.

Приложения дифференцирования

Еще Ньютон, вводя понятие производной, ставил при этом две основных задачи. Первая из них состояла в том, чтобы по заданной переменной (в терминологии Ньютона – *флюэнте*) найти ее производную (по Ньютону – *флюксию*), а вторая – в восстановлении переменной (т.е. функции) по известному значению ее производной. Первая задача относится к дифференциальному исчислению, а вот вторая – к процедуре, обратной к дифференцированию, т.е. к *интегрированию*. В механической интерпретации: в первом случае, по известному закону движения тела $x = x(t)$ определяется его скорость $v = x'$: во втором случае, по известной скорости v движущегося тела требуется восстановить закон его движения $x = x(t)$. В последнем случае искомая величина удовлетворяет соотношению (*уравнению*) $x' = v$. В математике соотношение в котором неизвестная функция входит под знаком производной называется *дифференциальным уравнением*.

Актуальность таких задач обусловлена, в первую очередь, естественными приложениями из механики. В частности, в основе динамики в значительной степени лежит *второй закон Ньютона.* Согласно ему ускорение a движущегося тела прямо пропорционально действующей на него силе F, *т.е.* справедливо знаменитое равенство $F = ma$, где m есть масса тела. Мы знаем, что скорость движущегося тела v представляет собой производную от его координаты x. Ну а ускорение a, характеризующее скорость изменения скорости v, оказывается тем самым производной v' от этой скорости или *второй производной* (производной от производной) x'' от координаты x. Тем самым второй закон Ньютона может быть записан в дифференциальной форме

$F = mx''$. На практике часто возникает задача определения закона движения тела $x = x(t)$ с известной массой m под действием заданной силы F. Таким образом, требуется найти такую функцию $x = x(t)$, которая удовлетворяет дифференциальному уравнению $mx'' = F$. В дальнейшем подобные задачи появились в химии (при изменении концентраций реагирующих веществ под действием химической реакции), биологии (при изменение численности биологических видов при тех или иных формах их сосуществования), экономике (при изменение ситуации на рынке в ходе различных экономических отношениях) и др.

Введение понятия производной привело к появлению одного из ведущих направлений современной математики – теории *дифференциальных уравнений*. Первые шаги в этом направлении были сделаны естественно творцами дифференциального исчисления – Ньютоном и Лейбницем. Их работы были продолжены учеником Лейбница Иоганном Бернулли, учеником последнего – Леонардом Эйлером и далее Жозефом Луи Лагранжем.

Параллельно развивалось другое приложение дифференциального исчисления. Как известно, производная характеризует скорость изменения функции: чем больше значение производной, тем быстрее изменяется функция. Еще Николя Орем, а позднее Иоганн Кеплер обратили внимание на то, что в окрестности точки *экстремума* (минимума или максимума) скорость изменения функции практически сходит на нет (рис. 6).

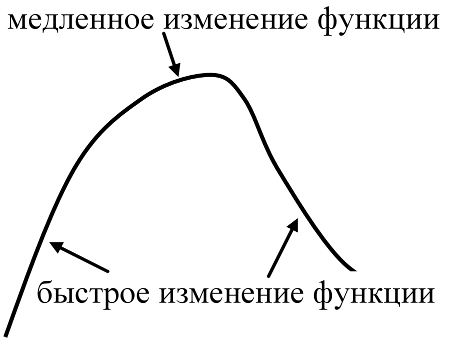

Рис. 6. Производная и экстремум.

Одним из непосредственных предшественников Ньютона и Лейбница был блистательный французский математик **Пьер Ферма** (рис. 7), имевший более или менее ясное представление о природе операции дифференцирования. Он утверждал, что в точке экстремума t производная функции x обращается в нуль. Таким образом, для того, чтобы найти экстремум заданной функции $x = x(t)$ требуется вычислить ее производную $x' = x'(t)$. Те точки t, в которых эта производная равна нулю (и только они!) могут оказаться точками экстремума данной функции. Тем самым поиск экстремума сводится к процедуре вычисления производной и решению уравнения $x'(t) = 0$. Поскольку искомой величиной здесь являются точка экстремума (число) t, то уравнение это оказывается *алгебраическим*. Пусть, к примеру, требуется узнать, в какой точке квадратичная функция $x(t) = t^2$ (т.е. парабола) принимает свое наименьшее значение. Следуя Ферма, мы сначала вычисляем производную этой функции $x'(t) = 2t$, а затем приравниваем ее нулю. В результате получаемое уравнение $2t = 0$ имеет решение $t = 0$, которое и является искомой точкой минимума параболы. Итак, отталкиваясь от понятия производной, Ферма свел задачу отыскивания минимумов и максимумов функций к решению алгебраических уравнений.

Рис. 7. Пьер Ферма (1601 – 1665)

Естественно, у Ферма еще нет строгого определения понятия производной, да и результат свой он получил исключительно для многочленов. Идущий же следом Лейбниц показал, что утверждение Ферма остается в силе для любой дифференцируемой функции. Он обратил внимание на то, что тип экстремума (минимум или максимум) определяется знаком второй производной в рассматриваемой точке. Дальнейшее развитие теории экстремума шло в значительной степени как приложение теории дифференцирования. Эстафету от Лейбница приняли его ученики – братья Якоб и Иоганн Бернулли, а далее Эйлер и Лагранж.

Теория дифференцирования нашла глубокие приложения в геометрии. Пусть, к примеру, имеется некоторая кривая, на которой выбраны точки A и B (рис. 8). Обозначим через α угол между касательными к кривой в указанных точках, а через s – длину дуги кривой, соединяющей эти точки. Тогда предел отношения α / s, когда точка B стремится к A, называется *кривизной* данной кривой в точке A. Связанное с касательными, а значит, с операцией дифференцирования понятие кривизны является количественной характеристикой кривой, позволяющей судить, если можно так выразиться, в какой степени «кривой» оказывается рассматриваемая кривая в окрестности заданной точки. Например, равенство кривизны нулю означает, что мы имеем дел с прямой. Если величина кривизны возрастает, то данная кривая становится все «кривее».

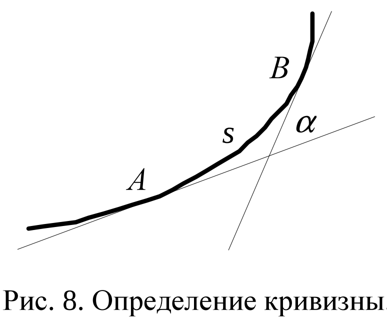

Рис. 8. Определение кривизны.

Применение аппарата дифференциального исчисления к геометрии привело к возникновению *дифференциальной геометрии*, ставшей одним из ведущих геометрических направлений. Леонард Эйлер, Гаспар Монж, Карл Фридрих Гаусс заложили основы дифференциальной геометрии кривых и поверхностей. Существенный прорыв в этом направлении сделал в середине 19 века великий немецкий математик **Бернхард Риман** (рис. 9). Согласно хорошо известному *пятому постулату Евклида*, через точку, взятую вне прямой, проходит ровно одна прямая, параллельная данной. Однако в *геометрии Лобачевского* таких прямых оказывается бесконечное множество, а в *геометрии Римана* вообще не существует параллельных прямых. Риман установил, что евклидова геометрия реализуется на поверхности нулевой кривизны, геометрия Лобачевского – на поверхности отрицательной кривизны, а его геометрии соответствует уже положительная кривизна. Риман, вводя аппарат дифференциального исчисления в геометрию, смог взглянуть на различные геометрические теории с единых позиций.

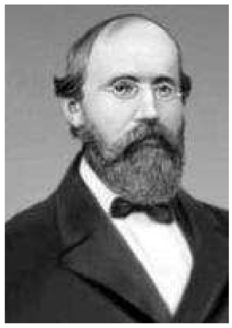

Рис. 9. Бернхард Риман (1826 –1866)

Постепенно совершенствовалась и сама теория дифференцирования. При многократном применении операции дифференцирования было получено представление о *старших производных*. Для функций многих переменных естественным образом определили *частные производные*. К примеру, для функции $f = f(x, y)$ можно определить ее частную производную по первому аргументу $\dfrac{\partial f}{\partial x}$, которая вводится как предел отношения $\left[f(x+\tau, y) - f(x, y) \right] / \tau$ при значении τ, стремящимся к нулю. Аналогичным образом определяется частная производная по второму аргументу, а также соответствующие старшие производные. Поскольку функции многих переменных широко встречаются на практике (характеристики системы могут меняться не только со временем, но и оказаться разными в различных точках пространства), *дифференциальные уравнения с частными производными* оказались важнейшим классом математических моделей процессов природы.

Свои особенности имеет дифференцирование *функций комплексного переменного*. Различные обобщения производной имеются в алгебре, топологии, геометрии. К примеру, в дифференциальной геометрии и связанным с ним *тензорным анализом* возникает понятие *ковариантной производной,* находящее широкое применение в теоретической физике. Физические приложения в значительной степени и предопределили дальнейший ход развития теории дифференцирования.

Обобщенное дифференцирование

Двадцатый век стал, в значительной степени, веком физики. Решение всё более сложных практических задач требовало разработки соответствующего математического аппарата. Выдающийся советский математик **Сергей Львович Соболев** (рис. 10) занимался исследованием *ударных волн* – поверхностей, при переходе через которые характеристики процесса (давление, скорость, плотность и др.) меняются скачком, и получалась какая-то странная ситуация. Согласно известным физическим законам рассматриваемый процесс описывался некоторыми дифференциальными уравнениями (точнее, уравнениями с частными производными). При этом, функции состояния этого процесса (относительно которых мы имеем соответствующие уравнения), исходя из физического смысла задачи, оказывались разрывными. Однако, производная от разрывной функции заведомо не имеет смысла. Тогда как же соотнести дифференциальные уравнения, выведенные на основе законов физики, с разрывным характером их решений, следующим из той же физики? В результате Соболев задается целью обобщить понятие дифференцирования на заведомо не дифференцируемые функции – негладкие, разрывные и т.д.

Рис. 10. Сергей Львович Соболев (1908 – 1989)

Как уже отмечалось, обратной операцией к дифференцированию является *интегрирование*. С геометрической точки зрения *интеграл* от функции $x = x(t)$ на отрезке $[a,b]$ численно равен площади S криволинейной фигуры, ограниченной координатной осью t, прямыми $t = a$ и $t = b$, а также рассматриваемой кривой. Указанная величина обозначается следующим образом:

$$S = \int_a^b x(t)dt.$$

Связь между операциями дифференцирования и интегрирования дает *формула интегрирования по частям*

$$\int_a^b x'(t)y(t)dt = -\int_a^b x(t)y'(t)dt + x(b)y(b) - x(a)y(a),$$

называемая *формулой Ньютона – Лейбница.* В частности, если хотя бы одна из рассматриваемых функций равна нулю на границах данного отрезка, то равенство принимает вид

$$\int_a^b x'(t)y(t)dt = -\int_a^b x(t)y'(t)dt.$$

От этого соотношения отталкивался Соболев.

Пусть мы хотим продифференцировать функцию $x = x(t)$, определенную на отрезке $[a,b]$ и обладающую весьма плохими свойствами. Наряду с ней, возьмем функцию $y = y(t)$, дифференцируемую сколь угодно большое число раз и обращающуюся в нуль на границах заданного отрезка. Тогда *обоб-*

щенной производной от функции x называется такой объект Dx, который удовлетворяет равенству

$$\int_a^b Dx(t)\,y(t)\,dt = -\int_a^b x(t)\,y'(t)\,dt$$

для любой функции y с указанным набором свойств. Его правая часть содержит интеграл от произведения заданной функции x на обычную производную y' от «хорошей» функции y. Тем самым под интегралом в правой части равенства стоит вполне осмысленная величина. Но, коль скоро мы имеем равенство с заданной правой частью, то этим определена и левая часть равенства, причем под обобщенной производной понимается тот самый объект, который находится там под интегралом и умножается на y.

Для дифференцируемых функций обобщенная производная совпадает с обычной. Но вот что произойдет, если мы попытаемся определить обобщенную производную от какой-либо «плохой» функции? Рассмотрим, к примеру, модуль, т.е. функцию $f(t)=|t|$ (рис. 11).

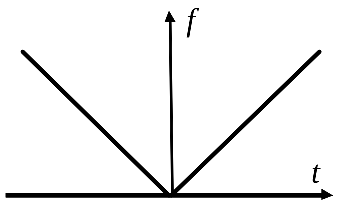

Рис. 11. Модуль.

Попробуем найти ее производную в нуле. Очевидно, отношение $\left[f(\tau)-f(0)\right]/\tau$ при положительных значениях τ равно 1, а при отрицательных значениях получаем -1. Таким образом, в результате перехода к пределу при τ, стремящимся к нулю, мы можем получить любое из значений +1 или -1, а не что-то конкретное. Более того, полагая $\tau=(-1)^k/k$ и устремляя параметр k в бесконечность, получаем, что τ стремится к нулю. Вот только предел указанного отношения в этом случае вообще не существует, поскольку мы имеем дело с последовательностью, элементы которой попеременно принимают значения +1 и -1. Это говорит о том, что модуль не имеет производной в нуле. Попробуем теперь найти обобщенную производную этой функции, рассматривая ее на отрезке $[-1,1]$.

Согласно приведенной выше формуле справедливо следующее равенство

$$\int_{-1}^{1} Df(t)y(t)dt = -\int_{-1}^{1} |t|\,y'(t)dt.$$

Учитывая, что модуль равен -1 для отрицательных значений t и 1 для положительных t, разобьем интеграл в последней формуле на две части – с интегрированием от -1 до 0 и от 0 до 1. Применяем приведенную ранее формулу интегрирования по частям к каждому из получаемых интегралов для функции $x(t)=t$. В результате имеем равенства

$$\int_{-1}^{0} x(t)y'(t)dt =$$

$$= -\int_{-1}^{0} x'(t)y(t)dt + x(0)y(0) - x(-1)y(-1),$$

$$\int_{0}^{1} x(t)y'(t)dt = -\int_{0}^{1} x'(t)y(t)dt + x(1)y(1) - x(0)y(0).$$

Производная x' от указанной функции x равна единице. Функция y по условию обращается в нуль на границах отрезка $[-1,1]$, а функция $x(t)=t$ равна нулю при $t=0$. В результате, формула, определяющая обобщенную производную от модуля, принимает вид

$$\int_{-1}^{1} Df(t)y(t)dt =$$

$$= -\int_{-1}^{0} y(t)dt + \int_{0}^{1} y(t)dt = \int_{-1}^{1} g(t)y(t)dt,$$

где функция g принимает значение -1 для отрицательных значений аргумента и +1 при их положительных значениях.

Сравнивая, левую и правую часть последнего равенства, справедливого для любой функции y, заключаем, что обобщенной производной Df от недифференцируемой функции f (модуля) является разрывная функция g.

Характерно, что она сохраняет естественные свойства обычной производной. В частности, для отрицательных значений аргумента функция f убывает с постоянной скоростью, а ее обобщенная производная (скорость изменения функции) g принимает постоянное отрицательное значение -1. При положительных значениях t функция f возрастает с постоянной скоростью, а ее обобщенная производная принимает постоянное положительное значение 1. Наконец, в нуле данная функция имеет излом, т.е. в окрестности нуля происходит ее качественное изменение с возрастания на убывание. Соответственно, ее обобщенная производная в нуле имеет разрыв, т.е. скорость изменения функции здесь изменилась скачком.

На основе описанной техники можно дифференцировать и разрывные функции. К примеру, обобщенная производная от той функции, что получилась в процессе дифференцирования модуля, характеризуется *δ-функцией Дирака*. Это и не функция даже, а существенно более общий объект, появившийся в квантовой механике и относящийся к классу *обобщенных функций*. Но и δ-функцию можно дифференцировать по схеме Соболева, как и любую обобщенную функцию…

Решения дифференциальных уравнений с производными, интерпретируемыми в обобщенном смысле, были названы *обобщенными решениями*. Работы в этом направлении были продолжены французским математиком Лораном Шварцем и другими авторами и стали ведущим направлением современной математической физики.

Таким образом, потребности физики стимулировали совершенствование представления математиков об операции дифференцирования.

Дифференцирование функционалов и операторов

Одновременно качественное обобщение понятия производной происходило с иной стороны, исходя из потребностей теории экстремума. Как отмечалось, Ферма установил, что необходимым условием экстремума функции в данной точке является равенство нулю его производной в этой точке. Что будет, если ищется экстремум не функции, а *функционала* – преобразования, действующего не на числа, а, например, на функции? Задачи такого типа возникают в механике.

К примеру, затраты энергии I движущегося тела на некотором интервале времени определяются обычно интегралом по этому интервалу от величины, зависящей от закона движения тела $x = x(t)$ и его производной (кинетическая энергия зависит от скорости, которая и является производной). Естественно поставить задачу нахождения такой функции $x = x(t)$, чтобы соответствующие ей затраты энергии были минимальными. Очевидно, каждому закону движения x соответствует конкретное значение затрат энергии $I = I(x)$. Тем самым величина I, зависящая от функции x, оказывается функционалом. С отдельными частными задачами такой природы сталкивались еще Ньютон и Лейбниц, а также братья Бернулли. Общий метод решения подобных задач впервые был разработан Эйлером. Однако более продуктивным оказался подход, предложенный великим французским математиком **Жозефом Луи Лагранжем** (рис. 12).

Рис. 12. Жозеф Луи Лагранж (1736 – 1813)

Можно, конечно, поступить испытанным способом, вычислив приращение функционала $I(x+\tau) - I(x)$. Однако что получится, если мы разделим это значение (т.е. число) на приращение τ, которое в данном случае само является функцией? Видимо, какая-то функция с не совсем понятными свойствами… И Лагранж предлагает вычислять приращение функционала в несколько иной форме, - $I(x+\sigma h)-I(x)$, где σ есть обычное число, а *h* – функция, аналогичная *x*. Эта величина затем делится на σ, благо результатом деления приращения функционала, т.е. числа, на другое число будет опять-таки число. Так вот, предел величины $\left[I(x+\sigma h)-I(x)\right]/\sigma$ при σ, стремящемся к нулю (если таковой существует), Лагранж назвал *вариацией* $\delta I(x)$ функционала *I* на функции *x*. Далее он показывает, что необходимым условием того, что функционал *I* имеет в качестве экстремума функцию *x*, является равенство нулю его вариации. Тем самым задача поиска минимумов и максимумов функционалов сводится к решению уравнения $\delta I(x)=0$ относительно функции *x*, что очень напоминало результат, установленный ранее Ферма. Это стало началом нового раздела математики, называемого *вариационным исчислением*.

В начале 20 века французский математик **Рене Гато** решил распространить конструкции Лагранжа с интегральных функционалов на функционалы общего вида. Их аргументами являются объекты произвольной природы, лишь бы только их можно было складывать, умножать на число, а также была бы возможность переходить к пределу. Действительно, пусть имеется зависимость $I=I(x)$, сопоставляющая каждому объекту x число $I(x)$. Можно вновь рассмотреть величину $\left[I(x+\sigma h)-I(x)\right]/\sigma$, оказывающуюся обычным числом, где *h* есть объект той же природы, что и *x*, а σ– некоторое число.

Предположим, что существует предел *l* указанной величины при σ, стремящемся к нулю. Очевидно, число *l* будет зависеть от *x* и *h*, т.е. $l=l(x,h)$. Возможно, что эта зависимость окажется *линейной* относительно *h*. Это означает, что справедливо равенство $l=\varphi(h)$, где величина φ (своя для каждого *x*) обладает тем свойством, что $\varphi(h+g)=\varphi(h)+\varphi(g)$ для любых объектов *h* и *g*, а $\varphi(ah)=a\varphi(h)$ для любого объекта *h* и числа *a*. В этих условиях φ называют *производной Гато* функционала *I* на объекте *x* и обозначают через $I'(x)$ по аналогии с производной обычной функции. Кстати, применяя конструкцию Гато к функциям, мы обычную производную и получим. Необходимым условием экстремума произвольного функционала *I* на объекте *x* оказывается равенство $I'(x)=0$, которое превращается в известное условие Ферма, если *x* является числом.

Схему Гато можно перенести с функционалов на преобразования, связывающие объекты произвольной природы и называемые *операторами*. Опять же, предполагается, что рассматриваемые объекты можно складывать, умножать на число, а также для них имеет смысл понятие предела (множества такой природы называются *линейными топологическими пространствами*). Пусть имеется оператор *I*, связывающий два таких множества *X* и *Y*. Рассматриваются объекты *x* и *h* из бласти определения *X*, а также некоторое число σ. Тогда величина $\left[I(x+\sigma h)-I(x)\right]/\sigma$ будет принадлежать области значений *Y* оператора *I*. Предположим, существует предел *l* указанной величины при σ, стремящемся к нулю, который принадлежит множеству *Y* и зависит от *x* и *h*, т.е. $l=l(x,h)$. Может оказаться, что зависимость $l=\varphi(h)$ этого предела от *h* вновь окажется линейной. Тогда φ (свое для каждого *x*) называется *производной Гато* оператора *I* на *x* и обозначается опять-таки через $I'(x)$. Аргументом φ служит объект *h*, принадлежащий множеству *X*, а величина $\varphi(h)$ (т.е. полученный предел) принадлежит множеству *Y*. Тем самым производная оператора оказывается преобразованием тех же самых множеств, что и исходный оператор *I*, только преобразованием линейным.

При переходе от производной функции к производной оператора начинает проявляться связь процедуры дифференцирования со свойствами линейности. Эта связь становится более явной в другом определении производной оператора, данном земляком и современником Гато – **Морисом Фреше**. В отличие от Гато, Фреше отталкивался от понятия дифференциала. Ранее отмечалось, что дифференциалом функции $x = x(t)$ в некоторой точке t называется главная линейная часть приращения функции, т.е. первое слагаемое в правой части равенства $x(t+\tau) - x(t) = x'(t)\tau + b(\tau)$, где величина b имеет более высокий порядок малости, чем τ в том смысле, что отношение $b(\tau)/\tau$ стремится к нулю при τ, стремящемся к нулю. Подобно Гато, Фреше рассматривает общее преобразование I множеств X и Y (линейных топологических пространств). Так вот, если существует линейный непрерывный оператор φ, связывающий указанные выше множества и удовлетворяющий соотношению $I(x+h) - I(x) = \varphi(h) + b(h)$, где величина b имеет более высокий порядок малости, чем h, то оператор этот называется *производной Фреше* преобразования I на x и обозначается через $I'(x)$. Она обладает более сильными свойствами, чем производная Гато, но зато применима для более узкого класса решаемых задач.

Производные операторов нашли широкие приложения во многих разделах математики, одним из них стала вычислительная математика.

Дифференцирование и вычислительная математика

Связь дифференцирования со свойствами линейности отчетливо проявляется в *вычислительной математике*, предметом которой является разработка приближенных методов решения тех или иных задач, в частности, различных уравнений. Пусть, к примеру, требуется найти решение t уравнения $x(t) = 0$, т.е. точку пересечения кривой x и координатной оси. Для его приближенного решения можно воспользоваться *методом касательных*, восходящим к Ньютону. Этот метод сводится к последовательному нахождению нового приближения t_{k+1} решения уравнения по известному значению t_k на предшествующем шаге k согласно формуле

$$t_{k+1} = t_k - \left[x'(t_k)\right]^{-1} x(t_k).$$

Геометрически величина $t(x_k)$ представляет собой длину катета AB треугольника ABC (рис. 13). Второй катет AC есть не что иное, как разность $t_{k+1} - t_k$, взятая с противоположным знаком. Отношение этих величин представляет собой тангенс угла α наклона касательной к кривой x в точке t_k.

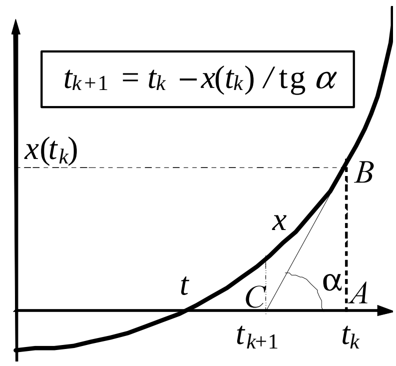

Рис. 13. Метод касательных.

Таким образом, метод касательных действует следующим образом. Пусть на предшествующем шаге алгоритма известно k-ое приближение решения уравнения. Тогда в точке t_k (точка A на координатной оси) проводится перпендикуляр к оси до пересечения с данной кривой. В получаемой при этом точке B на кривой строится касательная, которая пересекает координатную ось в некоторой точке C. Ее координата – это и есть новое приближение t_{k+1}. Далее процесс возобновляется, пока не будет найдена с желаемой степенью точности искомая точка t.

Выдающийся советский математик, лауреат нобелевской премии по экономике **Леонид Витальевич Канторович** (рис. 14). распространил метод касательных на уравнения общего вида. Действительно, пусть требуется решить уравнения $I(t) = 0$, где в качестве t может выступать не только число, но и вектор, функция, семейство функций и т.п., т.е. объект произвольной природы. При этом I представляет собой уже некоторый оператор. Формально алгоритм, называемый *методом Ньютона – Канторовича*, имеет тот же вид, что и описанный выше метод касательных. При известном *k*-ом приближении t_k новое приближение t_{k+1} определяется по формуле

$$t_{k+1} = t_k - \left[I'(t_k)\right]^{-1} I(t_k).$$

Рис. 14. Леонид Витальевич Канторович (1912 –1986)

Здесь величина $I'(t_k)$ представляет собой производную оператора *I* в точке t_k, которая, как известно, является линейным оператором. В обычном методе касательных величину $\left[x'(t_k)\right]^{-1}$ можно было понимать как единицу, деленную на число $x'(t_k)$, являющееся значением производной рассматриваемой функции в данной точке. Теперь же $\left[I'(t_k)\right]^{-1}$ выражает преобразование, обратное к $I'(t_k)$, т.е. такое, что последовательное выполнение преобразований $I'(t_k)$ и $\left[I'(t_k)\right]^{-1}$ любого объекта *y* дает этот же самый объект *y*. Таким образом, в результате действия оператором $I'(t_k)$ на указанное равенство мы получим эквивалентное ему соотношение

$$I'(t_k)(t_{k+1}) = I'(t_k)(t_k) - I(t_k).$$

В чем же смысл полученного результата? На *k*-ом шаге алгоритма мы знаем соответствующее приближение t_k, а значит, величину $I(t_k)$, а также производную $L = I'(t_k)$. Тем самым можно вычислить правую часть имеющегося равенства, которую обозначим через g. В результате относительно нового приближения t_{k+1} мы получаем уравнение $L(t_{k+1}) = g$. Спрашивается, что же мы выиграли, сведя старое уравнение к новому? Дело в том, что, будучи операторной производной, оператор *L* является линейным. Таким образом, метод Ньютона – Канторовича позволяет решение нелинейного уравнения $I(t) = 0$ свести к многократному решению линейных уравнений типа $L(t_{k+1}) = g$.

В частности, пусть *t* представляет собой вектор *n*-ого порядка, а соотношение $I(t) = 0$ является системой нелинейных алгебраических уравнений, решить которую весьма не просто. В этих условиях производная оператора *I* оказывается *матрицей,* называемой *якобианом,* а $L(t_{k+1}) = g$ представляет собой систему линейных алгебраических уравнений относительно вектора t_{k+1}. Естественно, решение (хотя бы и многократное) линейных уравнений представляется существенно более простой задачей, чем анализ уравнений нелинейных. Этим и объясняется высокая практическая значимость указанного подхода, тем более что математические модели процессов природы, как правило, приводят к нелинейным уравнениям. Ну, а решающим инструментом здесь оказывается операция дифференцирования, позволяющая свести нелинейную задачу к линейной.

Заключение. Дифференцирование и линеаризация

Всё происходящее в этом мире обусловлено какими-то причинами. Связывая причину и следствие на основе тех или иных законов природы, мы получаем *математическую модель* исследуемого процесса. В ряде случаев получаемый объект оказывается линейным, т.е. состояние системы в определенном смысле пропорционально входным факторам, влияющим на процесс. Тогда для анализа системы оказывается пригодным мощный алгебраический аппарат. Однако значительно чаще соотношение между входными и выходными факторами процесса оказывается все-таки нелинейным. *Нелинейные объекты* обладают существенно более сложной структурой и хуже поддаются анализу по сравнению с линейными. Тем не менее, существует общий принцип исследования нелинейных объектов.

Нелинейный анализ – это, прежде всего, *анализ локальный*. В малой области нелинейные свойства еще не успевают достаточно сильно проявиться. Тем самым в основе локального нелинейного анализа лежит *метод линеаризации*, позволяющий хотя бы частично распространить хорошо разработанный алгебраический аппарат на нелинейные объекты. В частности, гладкая кривая локально оказывается достаточно близкой к прямой – ее касательной, а гладкая поверхность – к соответствующей касательной плоскости. *Римановы геометрии*, обладающие той или иной кривизной, локально устроены так же, как евклидово пространство, имеющее нулевую кривизну. Одним из центральных понятий топологии является *гладкое многообразие*, которое локально устроено как линейное пространство. Во многих разделах математики широко применяется такой достаточно сложный объект, как *группа Ли*; однако локально она в значительной степени характеризуется своей алгеброй Ли, являющейся чисто алгебраическим объектом. Мощным инструментом *нелинейного функционального анализа* являются линеаризационные *теоремы о неявной и обратной функции*, позволяющие восстановить некоторые свойства нелинейного оператора по имеющимся свойствам его линейной аппроксимации. В *теории динамических систем* качественный анализ зачастую проводится на основе их линейного приближения. Важнейшим инструментом *теории экстремума* являются *необходимые условия оптимальности*, получаемые на основе линеаризации исследуемой системы. В вычислительной математике широко применяются линеаризационные алгоритмы типа описанного выше метода Ньютона – Канторовича, позволяющие численное решение нелинейной задачи сводить к последовательному решению семейства линейных задач.

Что же лежит в основе метода линеаризации? Основным аппаратом линеаризации и является рассматриваемая нами операция дифференцирования. В определении касательной прямой и касательной плоскости как раз и входит производная. Вспомним, что дифференциал – это главная линейная часть приращения функции. Соответственно, производная оператора есть оператор линейный.

Таким образом, именно в процессе дифференцирования мы получаем линейную аппроксимацию исследуемого объекта, и по мере постижения всё более и более сложных нелинейных явлений окружающего мира неминуемо будет и далее совершенствоваться аппарат дифференциального исчисления. В будущем, возможно, нас еще ожидают новые формы производных, которые позволят решить такие задачи, которые пока не поддаются анализу.

В заключение хочу выразить благодарность следующим специалистам, ознакомившимся с первоначальным вариантом статьи и сделавшим ряд важных замечаний – П.Г. Ицковой (Алматы), М.В. Ружанскому (Лондон), С.Ю. Симаранову (Москва), А.А. Симонову (Новосибирск), В.М. Сомсикову (Алматы) и Л.С. Сычевой (Новосибирск).

Литература: **[1.]** *Бурбаки Н.* Очерки по истории математики. – М., ИЛ, 1963. – 292 с.; **[2.]** *Вилейтнер Г.* История математики от Декарта до середины 19 столетия. – М., 1966. – 508 с.; **[3.]** *Клайн М.* Поиск истины. – М., Мир, 1988. – 296 с.; **[4.]** *Колмогоров А.Н.* Математика в ее историческом развитии. – М., Наука, 1991. – 224 с.;**[5.]** *Курант Р., Роббинс Г.* Что такое математика. – М., Просвещение, 1967. – 560 с.; **[6.]** *Панов В.Ф.* Математика древняя и юная. – М., МГТУ, 2006. – 648 с.; **[7.]** *Стиллвелл Дж.* Математика и ее история. – Москва, Ижевск, Институт космических исследований, 2004. – 530 с

Принято в печать 10.09.11

УДК 517.2

ЗАГАДКИ ДИФФЕРЕНЦИРОВАНИЯ

Семен Яковлевич Серовайский
Казахский национальный университет им. аль-Фараби, г. Алматы
Алматы 050078, пр. аль-Фараби 71
e-mail: serovajskys@mail.ru, тел. 275-39-34

Рассматривается эволюция взглядов на понятие производной.

RIDDLES OF THE DIFFERENTIATION

Simon Serovajsky
Al-Farabi Kazakh national University, Almaty
Almaty 050078, al-Farabi 71
e-mail: serovajskys@mail.ru, tel. 275-39-34

The evolution of the derivative conception is considered.

ДИФФЕРЕНЦИФЛДАУ ЖУМБАКТАРЫ

Семен Яковлевич Серовайский
Ал-Фараби Казак ултык унверситеті, Алматы к.
Алматы 050078, ал-Фараби дан. 71
e-mail: serovajskys@mail.ru, тел. 275-39-34

Туынды угымына пікірлер эволюциясы карастыралады.