Доклады Академии наук СССР
1957. Том 116, № 6

МАТЕМАТИКА

А. М. МОЛЧАНОВ

КОНЕЧНЫЕ МНОЖЕСТВА И СКАЛЯРНОЕ ПРОИЗВЕДЕНИЕ

(Представлено академиком М. В. Келдышем 8 V 1957)

Скалярное произведение обычно вводится как сумма произведений одноименных компонент векторов

$$(x, x') = x_1 x_1' + \ldots + x_s x_s'.$$

Цель настоящей заметки — показать, что скалярное произведение можно получить несложным предельным переходом, исходя из конечных множеств. Точнее, будет показано, что для подмножеств конечных множеств можно ввести понятие близости, переходящее в понятие скалярного произведения при увеличении числа элементов, входящих в конечное множество. Для того чтобы сформулировать точные утверждения, необходимо ввести несколько определений.

Определение 1. Разбиение конечного множества на непересекающиеся подмножества

$$M = m_1 + m_2 + \ldots + m_s$$

называется компонентой.

Подмножества m_1, m_2,..., m_s называются элементами компоненты.

Интерес представляет обычно случай большого числа компонент, так как здесь возникают предельные соотношения.

Определение 2. Зерном называется пересечение нескольких элементов различных компонент

$$p = m_{i_1}^{(1)} m_{i_2}^{(2)} \ldots m_{i_N}^{(N)}.$$

Определение 3. Две компоненты данной системы компонент («структуры») называются эквивалентными, если они состоят из одинакового числа множеств и соответственные зерна содержат одинаковое число элементов.

Например, если эквивалентны первая и вторая компоненты, то зерна $p = m_{i_1}^{(1)} m_{i_2}^{(2)} \ldots m_{i_N}^{(N)}$ и $\tilde{p} = m_{i_2}^{(1)} m_{i_1}^{(2)} \ldots m_{i_N}^{(N)}$, получающиеся друг из друга перестановкой i_1 и i_2, должны содержать одинаковое число элементов при всех $i_1, i_2, \ldots, i_N$.

Определение 4. Суперпозицией («смесью») двух зерен $p = m_{i_1}^{(1)} m_{i_2}^{(2)} \ldots m_{i_N}^{(N)}$ и $q = m_{k_1}^{(1)} m_{k_2}^{(2)} \ldots m_{k_N}^{(N)}$ называется зерно $r = m_{l_1}^{(1)} m_{l_2}^{(2)} \ldots m_{l_N}^{(N)}$, у которого каждый из множителей $m_{l_1}^{(l)}, m_{l_2}^{(2)}, \ldots, m_{l_N}^{(N)}$ совпадает с соответствующим множителем либо зерна p, либо зерна q.

Можно, конечно, говорить о суперпозиции не только двух, но и большего числа зерен.

Определение 5. Мерой близости («скалярным произведением») двух зерен p и q называется число совпадающих множителей

$$(p, q) = \delta_{i_1 k_1} + \delta_{i_2 k_2} + \ldots + \delta_{i_N k_N}.$$

Заметим, что мера близости двух зерен всегда не превосходит N, и если она равна N, то зерна тождественны.

Для удобства в дальнейшем будем пользоваться аддитивной записью зерна $m_{i_1}P_1 + m_{i_2}P_2 + \ldots + m_{i_N}P_N$ (вместо мультипликативной $m_{i_1}^{(1)} m_{i_2}^{(2)} \ldots$ $\ldots m_{i_N}^{(N)}$), введя «проекционные операторы» $P_1, P_2, \ldots, P_N$. Всюду в этой заметке предполагается, что все N компонент эквивалентны, а число элементов в каждой компоненте s фиксировано, в то время как число компонент N неограниченно увеличивается. Так как компоненты эквивалентны, то аддитивную запись можно упростить, собирая «подобные» при одинаковых (точнее, эквивалентных) множествах m: $p = m_{i_1}p_1 + m_{i_2}p_2 + \ldots$ $\ldots + m_{i_N}p_N = \sum_m mP(m)$. Таким образом зерно записывается как операторная функция над элементами компоненты $P(m)$. Мера близости двух зерен p и q удобно выражается через функции $P(m)$ и $Q(m)$: $(p, q) =$ $= \sum_m \mathrm{Sp}\,[P(m)\,Q(m)]$. Суперпозиция двух или нескольких зерен записывается, как легко проверить, при помощи проекционных операторов A_1, $A_2, \ldots, A_l$ (в сумме равных единичному) $P(m) = A_1P_1(m) + A_2P_2(m) + \ldots$ $\ldots + A_lP_l(m)$.

Для того чтобы получить (в пределе, конечно) обычное скалярное произведение в линейном векторном пространстве, рассмотрим множества, получающиеся объединением всех эквивалентных между собой (относительно перестановки компонент) зерен. Для краткости будем называть эти множества с и м м е т р и ч е с к и м и з е р н а м и. Ясно, что каждое симметрическое зерно определяется числовой функцией $n(m) = \mathrm{Sp}\ P(m)$, так как перестановка компонент — это то же, что перестановка операторов P_i. Симметрические зерна состоят, вообще говоря, из возрастающего вместе с N числа зерен. Если взять наугад из двух симметрических зерен по представителю и найти меру их близости, то ее значение будет, конечно, зависеть от того, какие были взяты представители. Однако при достаточно большом N в подавляющем числе случаев получатся весьма близкие результаты. При этом, если одно симметрическое зерно задается функцией $n(m)$, а другое функцией $n'(m)$, то мера близости их представителей в подавляющем большинстве случаев мало отличается от числа $\frac{1}{N}\,[n(m_1)\,n'(m_1) + n(m_2)\,n'(m_2) + \ldots + n(m_s)\,n'(m_s)]$. Таким образом, если ввести нормированные на единицу функции $a(m) = \frac{1}{N}\,n(m)$ и $a'(m) = \frac{1}{N}\,n'(m)$, а также нормированную на единицу меру близости $\frac{1}{N}\sum_m \mathrm{Sp}\,[P(m)\,P'(m)]$, то при $N \to \infty$ существует предел этой меры, и он равен обычному скалярному произведению функций $a(m)$ и $a'(m)$ (чем, собственно, и оправдывается само наименование «скалярное произведение» для меры близости зерен). Аналогично этому суперпозиция в пределе превращается в линейную комбинацию с суммой весов, равной единице. Поэтому на линейное пространство s-мерных векторов (или функций над s точками) с неотрицательными коэффициентами, в сумме равными единице, можно смотреть как на асимптотическое описание свойств структур — конечных множеств, разбитых на большое число эквивалентных компонент. Скалярное произведение возникает при этом естественным образом как асимптотика меры близости. (Заметим попутно, что сама мера близости имеет, по-видимому, аналог в теории информации, где вводится понятие близости двух сообщений. Однако неясно, соответствует ли чему-нибудь в теории информации рассмотрение эквивалентных компонент и симметрических зерен.)

Доказательство (и точная формулировка с остаточными членами) упомянутых выше предельных теорем получается сведением к теории сумматорных функций ([1]) и применением затем обычных теоретиковероятностных методов (характеристические функции и метод перевала). Проиллюстрируем это утверждение на примере скалярного произведения.

Так как скалярное произведение есть функция пары зерен, то основное множество, которое должно рассматриваться,— это множество последовательностей пар элементов $m_{i_1}^{(1)}, m_{k_1}^{(1)}; m_{i_2}^{(2)}, m_{k_2}^{(2)}; \ldots ; m_{i_N}^{(N)}, m_{k_N}^{(N)}$. Это множество есть прямое произведение N множеств вида (m, m'). На каждом из таких множеств зададим $2s$ функций следующим образом: $f_k(m, m') = \delta_{m_k m}$ и $f'_k(m, m') = \delta_{m'_k m'}$, $k = 1, 2, \ldots, s$. Рассмотрим далее $2s$ сумматорных функций $F_k = f_k^{(1)} + f_k^{(2)} + \ldots + f_k^{(N)}$ и $F'_k = f_k'^{(1)} + f_k'^{(2)} + \ldots + f_k'^{(N)}$. Скалярное произведение также является сумматорной функцией $G = g^{(1)} + g^{(2)} + \ldots + g^{(N)}$, причем в каждой компоненте $g = f_1 f'_1 + f_2 f'_2 + \ldots + f_s f'_s$. Ясно, что F_k не что иное, как «числа заполнения» множеств m_k, так что фиксирование значений F_k равносильно заданию симметрического зерна.

Следовательно, асимптотика скалярного произведения сводится к хорошо известной задаче теории сумматорных функций:

Заданы значения нескольких ($2s$) сумматорных функций $F_k = n(m_k) = Na(m_k)$, $F'_k = n'(m_k) = Na'(m_k)$, $k = 1, 2, \ldots, s$. Требуется найти среднее значение и дисперсию сумматорной функции G на множестве, определенном фиксированием значений F_k и F'_k.

Основную роль при решении такой задачи играет, как известно ([1]), структурная функция $\Omega_N(a_1, a_2, \ldots, a_s, a'_1, a'_2, \ldots, a'_s)$, равная числу элементов из M, на которых функции F_k и F'_k принимают заданные значения Na_k, Na'_k. Важное свойство, которым обладает Ω_N, состоит в том, что она есть свертка структурных функций, относящихся к отдельным компонентам. Это обстоятельство позволяет, введя характеристическую функцию — преобразование Лапласа структурной функции — применить, как это сделано в ([1]), аналитический аппарат теории вероятностей для доказательства предельных соотношений. Впрочем, нередко бывает удобно использовать прямо метод перевала. Недостаток места не позволяет остановиться более подробно на деталях рассуждений. По этой же причине приходится пройти мимо очень важного вопроса о том, каким образом могут появиться в качестве предельных образований линейные векторные пространства с действительными (а не только положительными) и, тем более, комплексными коэффициентами. Здесь можно только указать, что их появление связано с более глубоким изучением, как статистическим, так и структурным, конечных множеств, разбитых на эквивалентные компоненты.

Отделение прикладной математики
математического института им. В. А. Стеклова
Академии наук СССР

Поступило
7 V 1957

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] А. Я. Хинчин, Математические основания статистической механики, 1943.