Доклады Академии наук СССР
1973. Том 209, № 5

УДК 519.9+575.1 МАТЕМАТИКА

Ю. И. ЛЮБИЧ

К МАТЕМАТИЧЕСКОЙ ТЕОРИИ НАСЛЕДСТВЕННОСТИ

(Представлено академиком Ю. В. Линником 28 VI 1972)

Пусть F — квадратичное отображение n-мерного вещественного пространства $\mathbf{R}^n$ в себя:

$$x_j' = \sum_{i,\,k=1}^{n} a_{ik,\,j} x_i x_k, \quad j = 1, \dots, n, \tag{1}$$

причем $a_{ik,\,j} \geqslant 0$, $\sum\limits_j a_{ik,\,j} = 1$. Тогда симплекс σ^{n-1}: $x_i \geqslant 0$, $\sum x_i = 1$ инвариантен. Сужение отображения F на σ^{n-1} обозначим через V. В задачах математической генетики V задает эволюцию бесконечной «свободной популяции» и соответственно называется эволюционным оператором (см. [1]).

В работах С. Н. Бернштейна [2–5] была поставлена и для $n = 3$ решена проблема описания всех эволюционных операторов, удовлетворяющих «принципу стационарности»: $V^2 = V$. По мысли С. Н. Бернштейна, этому принципу должны удовлетворять все возможные элементарные механизмы биологической наследственности. Например, классическая менделевская наследственность в случае одного диаллельного локуса определяет оператор (закон Харди — Вайнберга [6, 7])

$$\begin{aligned} x_1' &= x_1^2 + x_1x_3 + {}^1/_4\, x_3^2, \\ x_2' &= x_2^2 + x_2x_3 + {}^1/_4\, x_3^2, \\ x_3' &= 2x_1x_2 + x_1x_3 + x_2x_3 + {}^1/_2\, x_3^2, \end{aligned} \tag{2}$$

для которого принцип стационарности верен.

Закон Харди — Вайнберга допускает параметризацию: $x_1' = p^2$, $x_2' = q^2$, $x_3' = 2pq$, где $p = x_1 + {}^1/_2 x_3$, $q = x_2 + {}^1/_2 x_3$. Параметры p, q связаны условиями $p \geqslant 0$, $q \geqslant 0$, $p + q = 1$ и интерпретируются как вероятности аллельных генов A, a в генофонде исходного поколения, состоящего из генотипов AA, aa, Aa с вероятностями x_1, x_2, x_3 соответственно. Принцип стационарности в этом случае эквивалентен законам сохранения генов $p' = p$, $q' = q$.

В работе [1] проблема Бернштейна была сужена до проблемы описания операторов, обладающих генной структурой. На фоне общей идеи генной структуры в [1] была определена так называемая элементарная генная структура (э.г.с.) и при дополнительном условии «нормальности» были описаны все операторы, обладающие э.г.с. (см. [1], теорема 5.3).

В настоящей работе после уточнения понятия генной структуры дается решение суженной проблемы Бернштейна.

Линейная форма $p(x)$, $x \in \mathbf{R}^n$, называется инвариантной, если $p(Vx) = p(x)$ на симплексе σ^{n-1}. Во всем пространстве это означает, что $p(Fx) = s(x)p(x)$, где $s(x) = \sum x_i$. Инвариантные линейные формы образуют подпространство J_V размерности $m \geqslant 1$. Рассмотрим в J_V многогранный телесный конус форм с неотрицательными коэффициентами.

Пусть $p_1, \ldots, p_N$ — базис этого конуса, нормированный в том смысле, что наибольший коэффициент каждой формы p_i равен единице. Впоследствии элементы этого базиса интерпретируются как вероятности генов.

Определение 1. Отображение μ: $\sigma^{n-1} \to \mathbf{R}^N$, задаваемое формулами $y_i = p_i(x)$, $i = 1, \ldots, N$, называется оператором мейоза.

Очевидно, $\mu V = \mu$ в силу инвариантности форм p_i.

Определение 2. Эволюционный оператор V обладает генной структурой (г.с.), если существует такое квадратичное отображение φ: $\mathbf{R}^N \to \sigma^{n-1}$ с неотрицательными коэффициентами, что $V = \varphi\mu$. При этом φ называется оператором оплодотворения.

Из наличия г.с. вытекает стационарность: $V^2 = \varphi\mu V = \varphi\mu = V$.

Случай э.г.с. выделяется требованием линейной независимости форм $p_1, \ldots, p_N$. Простейшим примером неэлементарной г.с. является открытый С. Н. Бернштейном «кадрильный» закон:

$$\begin{aligned} x_1' &= (x_1 + x_3)(x_1 + x_4), \quad x_2' = (x_2 + x_3)(x_2 + x_4), \\ x_3' &= (x_1 + x_3)(x_2 + x_3), \quad x_4' = (x_1 + x_4)(x_2 + x_4). \end{aligned} \tag{3}$$

Его генная структура была выявлена в ([1]). Там же было показано, что уже при $n = 3$ (т. е. в списке С. Н. Бернштейна) встречаются операторы, не имеющие г.с. ни в каком разумном смысле.

Укажем некоторое обобщение кадрильного закона. Пусть $n = kl$, $k > 1$, $l > 1$. Запишем n-мерный вектор x в виде матрицы:

$$x = \begin{pmatrix} x_1 & \ldots & x_l \\ x_{l+1} & \ldots & x_{2l} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ x_{(k-1)l+1} & \ldots & x_{kl} \end{pmatrix}. \tag{4}$$

Обозначим через $p_1, \ldots, p_k$ строчные, а через $q_1, \ldots, q_l$ — столбцевые суммы этой матрицы. Обобщенный кадрильный закон задается формулой $x' = p \otimes q$, т. е.

$$x' = \begin{pmatrix} p_1 q_1 & \ldots & p_1 q_l \\ p_2 q_1 & \ldots & p_2 q_l \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ p_k q_1 & \ldots & p_k q_l \end{pmatrix}. \tag{5}$$

При этом $x \mapsto (p_1, \ldots, p_k;\ q_1, \ldots, q_l)$ — оператор мейоза.

Оператор V называется приводимым, если существует такая пара координат, например, x_1, x_2, что x_j', $j = 1, \ldots, n$, зависят только от $x_1 + x_2$.

Оператор V называется вырожденным, если существует такая координата, например, x_1, что $x_1' = 0$.

Теорема. *Пусть эволюционный оператор V обладает генной структурой, неприводим и невырожден. Тогда это либо элементарная генная структура, либо обобщенный кадрильный закон.*

Доказательство основывается на следующем «принципе взаимности»: *если оператор обладает г.с., то для каждой координаты x_j, входящей в форму p_i, квадратичная форма x_j' делится на p_i.* На генетическом языке это означает, что если генотип j продуцирует в мейозе ген i, то этот ген обязан участвовать во всех оплодотворениях, производящих данный генотип.

Из принципа взаимности следует, что каждая координата x_j входит в одну или две из форм p_i. Координаты, входящие в одну форму, называются гомозиготными, остальные — гетерозиготными. Координата x_j называется чистой, если $a_{jj,j} = 1$, и расщепляющейся в противном случае. Для изучения этой альтернативы весьма полезна коммутативная алгебра $\mathfrak{A}_V$, определяемая в $\mathbf{R}^n$ оператором F: $xy = \Phi(x, y)$, где Φ — билинейное отображение, полярное к F. Аннулятор этой алгебры совпадает с подпространством $J_V^{\perp} = \{x: p(x) = 0 \quad \forall p \in J_V\}$ благодаря г.с. В фактор-алгебре $\mathfrak{A}_V / J_V^{\perp}$ умножение имеет вид $XY = {}^1/_2\{s(Y)X + s(X)Y\}$ (поскольку $s \in J_V$, она переносится на фактор-алгебру).

Стандартный базис $e_1, \ldots, e_n$ алгебры $\mathfrak{A}_V$ определяет в фактор-алгебре систему $E_1, \ldots, E_n$. Все E_i отличны от нуля и попарно различны. Система $\{E_i\}$ порождает многогранный конус. Оказывается, что крайним точкам этого конуса и только им соответствуют чистые координаты. Отсюда выводится, что все *гомозиготные координаты — чистые*. Гетерозиготные координаты могут быть как расщепляющимися, так и чистыми.

Гомозиготные и расщепляющиеся гетерозиготные координаты называются элементарными, чистые гетерозиготные — неэлементарными. *Доказывается, что либо все координаты элементарны, либо все неэлементарны.* В первом случае получается э.г.с., во втором — обобщенный кадрильный закон.

Условия неприводимости и невырожденности можно отбросить. Ответ усложнится, но останется обозримым. Во всяком случае он достаточно формально вытекает из основной теоремы.

Автор выражает благодарность Г. Р. Белицкому и А. А. Кириллову, обнаружившим некоторые дефекты первоначального варианта работы.

Харьковский государственный университет
им. А. М. Горького

Поступило
6 V 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Ю. И. Любич, УМН, **26**, в. 5, 51 (1971). [2] С. Н. Бернштейн, Наука на Украине, в. 1, 14 (1922). [3] S. N. Bernstein, C.R., **177**, 528 (1923). [4] S. N. Bernstein, C.R., **177**, 581 (1923). [5] С. Н. Бернштейн, Уч. зап. н.-и. кафедр Украины, отд. матем., в. 1, 83 (1924). [6] G. N. Hardy, Science, **28**, 49 (1908). [7] W. Weinberg, Jahresh. Ver. f. Vaterl. Naturk. Würt., **64**, 368 (1908).