

УДК: 519.7

Неполные системы линейных уравнений с ограничениями на переменные

И. Г. Минкевич

ФГБУН «Институт биохимии и физиологии микроорганизмов им. Г. К. Скрыбина РАН»,
Россия, 142290, г. Пущино, Московская обл., просп. Науки, д. 5

E-mail: minkevich@ibpm.pushchino.ru

Получено 16 июля 2014 г.,
после доработки 10 октября 2014 г.

Сформулирована задача описания объектов различной природы на основе системы линейных уравнений, в которой число неизвестных превосходит число уравнений. Важной особенностью такой задачи, существенно усложняющей ее решение, являются ограничения на значения ряда переменных. Примером такой задачи является выбор биохимических реакций, осуществляющих преобразование заданного субстрата (исходного вещества) в заданный продукт. В этом случае неизвестными являются скорости биохимических реакций, образующие искомым вектор решения. Компоненты этого вектора в описываемом подходе разделяются на две группы: 1) задаваемые, \vec{y} ; 2) зависящие от задаваемых, \vec{x} . Изучены варианты конфигурации области допустимых значений \vec{y} , следующие из ограничений, наложенных на компоненты \vec{x} . Выявлено, что часть ограничений могут быть излишними и поэтому исключенными из рассмотрения, что упрощает решение задачи. Анализируются случаи, когда два или более ограничений на \vec{x} приводят к появлению жестких связей между компонентами \vec{y} . Описаны методы поиска базисных решений, учитывающие особенности данной задачи. Постановка общей задачи и полученные решения проиллюстрированы биохимическим примером.

Ключевые слова: линейные уравнения, прямоугольные матрицы, линейные неравенства, стехиометрия метаболизма, метаболические пути

Incomplete systems of linear equations with restrictions of variable values

I. G. Minkevich

G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms RAS,
5 Nauki avenue, Pushchino, Moscow Region, 142290, Russia

Abstract. — The problem is formulated for description of objects having various natures which uses a system of linear equations with variable number exceeding the number of the equations. An important feature of this problem that substantially complicates its solving is the existing of restrictions imposed on a number of the variables. In particular, the choice of biochemical reaction aggregate that converts a preset substrate (a feedstock) into a preset product belongs to this kind of problems. In this case, unknown variables are the rates of biochemical reactions which form a vector to be determined. Components of this vector are subdivided into two groups: 1) the defined components, \vec{y} ; 2) those dependent on the defined ones, \vec{x} . Possible configurations of the domain of \vec{y} values permitted by restrictions imposed upon \vec{x} components have been studied. It has been found that a part of restrictions may be superfluous and, therefore, unnecessary for the problem solving. Situations are analyzed when two or more \vec{x} restrictions result in strict interconnections between \vec{y} components. Methods of search of the basis solutions which take into account the peculiarities of this problem are described. Statement of the general problem and properties of its solutions are illustrated using a biochemical example.

Keywords: linear equations, non-square matrices, linear inequalities, stoichiometry of metabolism, metabolic pathways

Citation: *Computer Research and Modeling*, 2014, vol. 6, no. 5, pp. 719–745 (Russian).

Введение

В ряде научно-исследовательских задач (относящихся, например, к стехиометрии клеточного метаболизма [Schilling et al., 1999; Papin, J. A., 2004] и к объектам иной природы [Васильев, 2011]) объект исследования описывается системой линейных уравнений, число которых меньше числа входящих в них переменных (см. пример, приведенный в следующем разделе). Физически это означает, что на эти переменные наложены также иные связи (дифференциальные уравнения и др.), описывающие регуляцию внутри объекта и (или) влияние внешних условий. Если эти связи не включены в общую систему уравнений, то описание объекта является неполным. Тем не менее ряд важных свойств таких объектов может быть найден с использованием только вышеупомянутых линейных уравнений. Привлекательность такого подхода основывается не только на значимости результатов, но и на том, что системы линейных уравнений позволяют найти общее решение для любой размерности.

С другой стороны, физические (в общем смысле этого слова) свойства объекта могут накладывать ограничения на диапазон допустимых значений переменных. В случае совокупности биохимических реакций такие ограничения основаны на том, что часть этих реакций необратима вследствие термодинамических свойств реагентов. Полное математическое описание, если оно сделано корректно, автоматически соблюдает такое требование (например, фазовые траектории не пересекают оси координат, за исключением некоторых стационарных точек). Если же математическая модель включает в себя только неполную систему линейных уравнений, то ее необходимо дополнить соответствующим числом неравенств. Это усложняет задачу. Если ее размерность не мала (больше, чем 3), то решение требует применения численных методов. Но и при этом, как будет видно ниже, процесс решения таких задач может далеко продвинуться в аналитическом виде, что в конечном счете заметно сокращает потребность в ресурсах для численных расчетов.

Данная работа посвящена исследованию общих свойств решений неполных систем линейных уравнений с ограничениями на допустимые значения переменных. При этом никакие требования типа экстремальности какой-либо оценочной функции не накладываются.

Пример, поясняющий постановку задачи

Рассмотрим пример из области стехиометрии метаболизма живых клеток. Для этого выберем из огромного количества биохимических реакций, протекающих в клетках, некоторое ограниченное их число, а именно реакции, которые могут быть вовлечены в какой-то процесс, имеющий физиологическое или биотехнологическое значение. Пусть это будут реакции, осуществляющие преобразование заданного субстрата в заданный продукт и одновременно обеспечивающие энергетические потребности клеток макроэргическими соединениями — АТФ и ГТФ — за счет процессов окисления, протекающих в электрон-транспортной цепи [Buchanan et al., 2000]. Считаем поток субстрата, входящий в эту совокупность реакций, и поток образуемого продукта известными.

Задача состоит в том, что нужно выбрать из рассматриваемого множества реакций те, что необходимы для осуществления заданного процесса, определить потоки через них и количества получаемых АТФ и ГТФ. Если необходимый набор реакций найден, то не составляет труда изобразить их в виде метаболической схемы, где они расположены в соответствующей последовательности. Заранее ясно, что если совокупность реакций выбрана недостаточной, то задача решения не имеет. С другой стороны, если эта совокупность больше минимально необходимой, то возможны различные варианты метаболических путей, а какие-то реакции могут вообще оказаться ненужными для осуществления данного процесса. Последние легко определяются — потоки через них равны нулю. Отсюда следует, что задача выбора реакций сводится к задаче нахождения совокупности потоков.

В качестве субстрата-источника энергии возьмем пируват (пировиноградную кислоту), в качестве продукта сукцинат (янтарную кислоту), в качестве реакций те, что входят в цикл трикарбонных кислот (цикл Кребса), некоторые соседние с ними реакции и электрон-транспортную цепь. Список этих реакций дан в таблице 1. Правый ее столбец содержит индексы направления реакций. Если индекс равен нулю, то реакция обратима. Если он равен 1, реакция необратима и разрешено только ее направление, указанное знаками стехиометрических коэффициентов (см. таблицу 2). Эти индексы имеют качественный смысл и в систему уравнений не входят. Таблица 2 содержит стехиометрическую матрицу v_{jk} [Степанов и др., 1976] и названия метаболитов, участвующих в рассматриваемых реакциях. Матрица v_{jk} в таблице 2 обведена жирной линией. Здесь j — номер метаболита, k — номер реакции. Каждый k -й столбец содержит стехиометрические коэффициенты соответствующей реакции — числа молекул, участвующих в ней. Отрицательные коэффициенты относятся к субстратам (потребляемым веществам), положительные — к продуктам (образуемым веществам). Если $v_{jk} = 0$, то j -е вещество в k -реакции не участвует. Нулевые элементы v_{jk} в таблице 2 затенены, чтобы было лучше видно элементы $v_{jk} \neq 0$.

Таблица 1. Биохимические реакции, выбранные для примера (см. текст)

k	Реакция	Индекс необратимости
1	Электрон-транспортная цепь (комплекс I)	0
2	Электрон-транспортная цепь (остальная часть)	1
3	Пируватдекарбоксилаза	1
4	Пируватдегидрогеназный комплекс	1
5	Пируваткарбоксилаза	1
6	Цитратсинтаза	1
7	Аконитаза (стадия 1)	0
8	Аконитаза (стадия 2)	0
9	Изоцитратдегидрогеназа (стадия 1)	1
10	Изоцитратдегидрогеназа (стадия 2)	1
11	α -кетоглутаратдегидрогеназа	1
12	Сукцинил-КоА-синтаза (ГТФ)	0
13	Сукцинатдегидрогеназа	0
14	Фумараза	0
15	Малатдегидрогеназа	0
16	Ацетальдегиддегидрогеназа	0
17	Изоцитратлиаза	0
18	Малатсинтаза	1
19	Алкогольдегидрогеназа	0
20	Лактатдегидрогеназа	0

Неизвестными в данной задаче являются потоки через реакции, z_k . Величины z_k измеряются в молях реакционных актов, так что количество j -го вещества, прореагировавшего в k -й реакции, равно $v_{jk} z_k$. В обратимых реакциях возможен любой знак z_k . В необратимых реакциях (индекс необратимости которых равен 1) отрицательные значения z_k запрещены. Заданное множество реакций может обмениваться с остальным метаболизмом или с внеклеточной средой потоками b_j . Величина b_j отрицательна, если j -е вещество в целом потребляется данной системой (неважно, одной или несколькими реакциями или происходит как потребление, так и образование, но в общем балансе это вещество потребляется), положительна, если оно обра-

зуется. Если же $b_j = 0$, то образование и потребление j -го метаболита сбалансированы в системе реакций и обмен этим веществом с остальным метаболизмом или с внеклеточной средой отсутствует. Баланс каждого j -го метаболита имеет вид

$$\sum_k v_{jk} z_k = b_j. \quad (1)$$

Таблица 2. Стехиометрическая матрица (пояснения в тексте)

j	i	метаболиты	$k \rightarrow$																			
\downarrow	\downarrow		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		O ₂	0	-0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2		CO ₂	0	0	1	1	-1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
3		H ₂ O	1	2.5	0	0	-1	-1	1	-1	0	0	0	0	-1	0	0	0	0	-1	0	0
4		H ⁺	-1	0	0	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	1	1
5		H ₃ PO ₄	-1	-1.5	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6		АТФ	1	1.5	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7		АДФ	-1	-1.5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8		ГТФ	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
9		ГДФ	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
10	1	НАДН	-1	0	0	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	1	1
11	2	НАД ⁺	1	0	0	-1	0	0	0	0	-1	0	-1	0	0	0	-1	-1	0	0	-1	-1
12	3	убихинол	1	-1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
13	4	убихинон	-1	1	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
14	5	КоА	0	0	0	-1	0	1	0	0	0	0	-1	-1	0	0	0	-1	0	1	0	0
15	6	ацетил-КоА	0	0	0	1	0	-1	0	0	0	0	0	0	0	0	0	1	0	-1	0	0
16	7	сукцинил-КоА	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
17	8	пируват	0	0	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
18	9	ацетальдегид	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	1	0
19	10	оксалоацетат	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
20	11	цитрат	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
21	12	цис-аконитат	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0
22	13	изоцитрат	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	-1	0	0	0
23	14	оксалосукцинат	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0
24	15	α-кетоглутарат	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0
25	16	сукцинат	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	1	0	0
26	17	фумарат	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0
27	18	L-малат	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	1	0	0
28	19	глиоксилат	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0
29	20	этанол	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
30	21	лактат	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1

В данном примере матрица v_{jk} имеет размер 30×20 , то есть число строк в ней больше числа столбцов. Однако по физическому смыслу задачи вся v_{jk} не может рассматриваться как матрица системы уравнений для неизвестных z_k . Для окисления субстрата требуется кислород, продуктами окисления являются двуокись углерода и вода; этими веществами клетки обмениваются с окружающей средой, и их количества должны быть вычислены, а не заданы. Кроме того, носители энергии АТФ и ГТФ образуются из АДФ, ГДФ и ортофосфата H₃PO₄; последние берутся из остального метаболизма, а АТФ и ГТФ отдаются в остальной метаболизм. Потоки b_j , соответствующие этим веществам, также заранее не известны и должны быть вычислены.

Таблица 3. Правые части уравнений (см. текст)

$j \downarrow$	$i \downarrow$	Метаболиты	b_j
1		O ₂	Должны быть вычислены
2		CO ₂	
3		H ₂ O	
4		H ⁺	
5		H ₃ PO ₄	
6		АТФ	
7		АДФ	
8		ГТФ	
9		ГДФ	
10	1	НАДН	0
11	2	НАД ⁺	0
12	3	убихинол	0
13	4	убихинон	0
14	5	КоА	0
15	6	ацетил-КоА	0
16	7	сукцинил-КоА	0
17	8	пируват	-2
18	9	ацетальдегид	0
19	10	оксалоацетат	0
20	11	цитрат	0
21	12	<i>цис</i> -аконитат	0
22	13	изоцитрат	0
23	14	оксалосукцинат	0
24	15	α -кетоглутарат	0
25	16	сукцинат	1
26	17	фумарат	0
27	18	L-малат	0
28	19	глиоксилат	0
29	20	этанол	0
30	21	лактат	0

Таким образом, данная задача делится на две части. Строки матрицы v_{jk} (таблица 2), расположенные ниже двойной разделительной линии, входят в первую часть задачи, в которой находятся потоки z_k . Строки, расположенные выше, — во вторую часть, где с использованием найденных z_k вычисляются неизвестные b_j . Заданные и вычисляемые потоки b_j аналогичным образом разделены в таблице 3. Нумерация компонент вектора b_j , в общем, произвольна. Поэтому вычисляемые компоненты b_j не обязательно должны быть собраны в один сплошной массив, они могут быть расположены вперемешку с известными b_j .

Назовем ту часть общей задачи, в которой правые части уравнений заданы, а z_k находятся, основной задачей. Остальную часть, в которой соответствующие b_j должны вычисляться, — присоединенной задачей. Очевидно, что главная проблема состоит в решении основной задачи, после чего решить присоединенную задачу не составляет труда.

Для строк v_{jk} в «нижней» ее части (то есть для метаболитов, входящих в основную задачу) введем отдельную нумерацию i . Тогда из (1) выделяется система линейных уравнений

$$\sum_k v_{ik} z_k = b_i. \quad (2)$$

В данном примере система (2) имеет размер 21×20 , в ней число строк пока еще превышает число столбцов. Однако ранг этой матрицы $\text{rank } v_{ik} = 18$. Исключение линейно зависимых строк оставляет от v_{ik} матрицу размером 18×20 , и, таким образом, задача приходит к названному во введении виду.

Отметим, что поддержание нулевых значений ряда компонент вектора b_i осуществляется кинетическими регуляторными механизмами клеток, действие которых в данной постановке задачи не рассматривается. Полагается, что это действие обеспечивает такую сбалансированность функционирования системы реакций, что и имеет место в действительности. Такое оставление биохимической кинетики «за кадром» обеспечивает линейность данной задачи.

Общие свойства систем линейных уравнений с прямоугольной матрицей и ограничениями на ряд неизвестных рассматриваются ниже. После нахождения z_k эти величины подставляются в остальные строки системы (1) (входящие в присоединенную задачу), откуда находят оставшиеся потоки b_j . Результаты, относящиеся к данному примеру, приведены ниже в разделе «Базисные решения биохимической задачи».

Общая постановка задачи

Исходная система уравнений для неизвестных переменных $\bar{z} = \{z_k\}$:

$$M\bar{z} = \bar{b}, \quad (3)$$

или детально:

$$\sum_{k=1}^K m_{ik} z_k = b_i, \text{ где } i = 1, 2, \dots, I, \text{ причем } I < K. \quad (4)$$

Далее полагаем, что матрица этой системы $M = \{m_{ik}\}$ обладает следующими двумя свойствами. 1. M — матрица полного ранга; он равен ее меньшему (вертикальному) размеру: $\text{rank } M = I$. Если ранг M меньше, чем I , то этому требованию можно удовлетворить, исключая уравнения, являющиеся линейными комбинациями остальных (безразлично, какие уравнения из комбинаций, равных нулю, выбрасываются). 2. Линейно независимыми друг от друга являются первые I столбцов матрицы M . Этого можно добиться, перенумеровывая столбцы M и соответственно переменные z_k .

На часть переменных z_k наложены ограничения, а именно: эти переменные должны быть неотрицательны. Обозначая соответствующие номера через k_r , имеем

$$z_{k_r} \geq 0. \quad (5)$$

Эти условия не снижают общность постановки задачи. Если физические свойства объекта таковы, что для каких-то k_r ограничение должно иметь вид $z_{k_r} \leq 0$, то, меняя для этих k_r знаки z_{k_r} и m_{ik_r} (последние для всех i) на противоположные, мы преобразовываем задачу к вышеописанному виду. Аналогично, если какие-то ограничения имеют вид $z_{k_r} \geq R_{k_r}$, то преобразование будет сделано введением новых переменных $z'_{k_r} = z_{k_r} - R_{k_r}$; это приведет к изменению величин b_i .

Если бы матрица M была квадратной ($I = K$), а условия (5) отсутствовали, то решением системы (3) было бы единственное значение вектора \bar{z} . Это решение нульмерно, так как является точкой в пространстве $\{z_k\}$. Наличие условий (5) в этом случае разрешает существование

такого решения, если ни одна компонента \bar{z} с номерами k_r не противоречит (5). Тот факт, что число столбцов M больше числа строк, существенно меняет свойства решения, выявлению которых посвящено дальнейшее изложение.

Преобразование задачи

Матрица M может быть представлена как состоящая из двух матриц:

$$M = (A \ C), \quad (6)$$

где A — квадратная $(I \times I)$ матрица, включающая в себя первые I столбцов исходной матрицы M , а C составлена из остальных столбцов M . Представление (6) подразумевает, что сплошная нумерация k может быть заменена на раздельную: $n=1,2,\dots,I$ для столбцов A и $l=1,2,\dots,(K-I)$ для столбцов C . Отсюда $A = \{a_{in}\}$, $C = \{c_{il}\}$. Соответственно, вектор неизвестных величин \bar{z} представляется как состоящий из двух векторов $\bar{z} = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$, где компоненты

$\bar{x} = \{x_n\}$ и $\bar{y} = \{y_l\}$ занумерованы вышеупомянутыми индексами n и l . Тогда исходная система уравнений может быть представлена в виде

$$A\bar{x} = \bar{b} - C\bar{y}, \quad (7)$$

или, в детальной форме,

$$\sum_{n=1}^I a_{in} x_n = b_i - \sum_{l=1}^{K-I} c_{il} y_l, \text{ где } i=1,2,\dots,I. \quad (8)$$

Ограничения (5) приобретают вид

$$x_{n_r} \geq 0, \quad y_{l_r} \geq 0 \quad (9)$$

где номера n_r и l_r соответствуют номерам k_r из первоначальной нумерации компонент вектора \bar{z} .

Решение в отсутствие ограничений на переменные

Оставим пока в стороне вопрос о том, удовлетворяет ли решение системы линейных уравнений ограничениям (9).

Обозначим правую часть системы (7) как $\bar{w} = \bar{b} - C\bar{y}$. Согласно свойству $2 \operatorname{rank} A = \operatorname{rank} M = I$. Поэтому если \bar{w} задано, то решение системы (7) находится однозначно. Поскольку матрицы A и C , а также вектор \bar{b} имеют известные фиксированные значения, то требуется задать вектор \bar{y} , то есть $K-I$ его компонент. Тогда

$$\bar{x} = A^{-1}\bar{w}, \quad (10)$$

где A^{-1} — матрица, обратная A . Другой вид выражения (10), который понадобится ниже:

$$\bar{x} = A^{-1}\bar{b} - A^{-1}C\bar{y} = \bar{f} - H\bar{y}, \quad (11)$$

где $\bar{f} = A^{-1}\bar{b}$, $H = A^{-1}C$. В детальной форме (11) имеет вид

$$x_n = \sum_{i=1}^I (a^{-1})_{ni} b_i - \sum_{l=1}^{K-I} \left(\sum_{i=1}^I (a^{-1})_{ni} c_{il} \right) y_l = f_n - \sum_{l=1}^{K-I} h_{nl} y_l, \quad (12)$$

где $L = \dim \{y_i\} = K - I$ — размерность пространства $\{y_i\}$ (то есть число задаваемых компонент полного вектора решения \vec{z}), $H = \{h_{nl}\}$. Вместе векторы \vec{x} и \vec{y} образуют решение исходной системы $\vec{z} = \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix}$.

Задать \vec{y} можно бесконечным числом способов. Однако поскольку рассматриваемая здесь часть общей задачи линейна, то в пространстве $\{y_i\}$ можно выбрать базис из линейно независимых векторов $\vec{y}^{(\alpha)}$. Тогда любой вектор \vec{y} может быть представлен в виде

$$\vec{y} = \sum_{\alpha} Y^{(\alpha)} \vec{y}^{(\alpha)}, \quad (13)$$

где α — номер каждого из *этих* базисных векторов. Очевидно, что общее число линейно независимых $\vec{y}^{(\alpha)}$ равно размерности пространства $\{y_i\}$, то есть $L = K - I$. Следовательно, представление (13) является полным при $\alpha = 1 \div L$. Для каждого $\vec{y}^{(\alpha)}$ правая часть системы (7) равна $\vec{w}^{(\alpha)} = \vec{b} - C\vec{y}^{(\alpha)}$, и тогда \vec{x} имеет частное значение $\vec{x}^{(\alpha)}$:

$$\vec{x}^{(\alpha)} = A^{-1}\vec{w}^{(\alpha)}. \quad (14)$$

Полностью частное решение исходной системы равно $\vec{z}^{(\alpha)} = \begin{pmatrix} \vec{x}^{(\alpha)} \\ \vec{y}^{(\alpha)} \end{pmatrix}$.

Задача, однако, заключается в том, чтобы найти базис не для \vec{y} (поскольку это лишь часть компонент искомого решения \vec{z}), а для общего выражения именно \vec{z} . Тогда общее решение для \vec{x} следует представить как разложение по $\vec{w}^{(\alpha)}$:

$$\vec{x} = \sum_{\alpha} X^{(\alpha)} \vec{x}^{(\alpha)} = \sum_{\alpha} X^{(\alpha)} A^{-1} \vec{w}^{(\alpha)} = \sum_{\alpha} X^{(\alpha)} A^{-1} (\vec{b} - C\vec{y}^{(\alpha)}) = \sum_{\alpha} X^{(\alpha)} (\vec{f} - H\vec{y}^{(\alpha)}). \quad (15)$$

Каждое частное значение $\vec{w}^{(\alpha)} = -C\vec{y}^{(\alpha)} + \vec{b}$ является линейной комбинацией из L столбцов матрицы $-C = -c_{il}$ и вектор-столбца \vec{b} , то есть из $L + 1$ векторов, причем вектор-столбцы $-c_{il}$ входят в эту комбинацию с коэффициентами y_i , а \vec{b} — с коэффициентом 1. Для получения полного базиса, в котором могут быть представлены все решения \vec{x} , нужно в общем случае задать $L + 1$ линейно независимых векторов $\begin{pmatrix} \vec{y}^{(\alpha)} \\ 1 \end{pmatrix}$, каждый из которых имеет $L + 1$ компоненту, одна из которых фиксирована. В особом случае, если ранг расширенной матрицы $(-C \quad \vec{b})$ меньше, чем $L + 1$, коэффициенты $y_i^{(\alpha)}$ могут давать значения $\vec{w}^{(\alpha)}$, линейно зависящие между собой, вследствие чего будут получаться линейно зависящие друг от друга решения системы (14) $\vec{x}^{(\alpha)}$. Однако поскольку для каждого $\vec{w}^{(\alpha)}$ необходимо найти полное решение

$\vec{z}^{(\alpha)} = \begin{pmatrix} \vec{x}^{(\alpha)} \\ \vec{y}^{(\alpha)} \end{pmatrix}$, то его часть, равная $\vec{y}^{(\alpha)}$, будет обеспечивать линейную независимость всех базисных решений $\vec{z}^{(\alpha)}$. Следовательно, в любом случае необходимо задавать $L + 1$ линейно независимых векторов $\begin{pmatrix} \vec{y}^{(\alpha)} \\ 1 \end{pmatrix}$.

Система (14) в отсутствие ограничений на переменные будет иметь $L + 1$ базисных решений, и α в (15) должно иметь диапазон $\alpha = 1 \div (L + 1)$. Полное решение исходной задачи имеет вид

$$\bar{z} = \sum_{\alpha=1}^{L+1} X^{(\alpha)} \bar{z}^{(\alpha)} = \sum_{\alpha=1}^{L+1} X^{(\alpha)} \begin{pmatrix} \bar{x}^{(\alpha)} \\ \bar{y}^{(\alpha)} \end{pmatrix}. \quad (16)$$

Поскольку \bar{y} — задаваемый вектор, он может быть представлен через $\bar{y}^{(\alpha)}$ с любыми коэффициентами, в том числе и $X^{(\alpha)}$.

Таким образом, общее решение системы линейных уравнений с числом уравнений, меньшим числа неизвестных, имеет $L + 1 = K - I + 1$ степеней свободы. Для его нахождения нужно задать $L + 1$ линейно независимых векторов $\begin{pmatrix} \bar{y}^{(\alpha)} \\ 1 \end{pmatrix}$ и вычислить $L + 1$ частных решений (14).

Ограничения на задаваемые переменные

Ограничения (9) относятся к обеим группам компонент общего решения \bar{z} , но поскольку вектор \bar{x} есть функция от задаваемого вектора \bar{y} , то в конечном счете все ограничения должны быть выражены в виде неравенств, относящихся только к \bar{y} . Неравенства $y_i \geq 0$ относятся непосредственно к \bar{y} . Неравенства $x_{n_r} \geq 0$, согласно (12), имеют вид

$$x_{n_r} = f_{n_r} - \sum_{l=1}^L h_{n_r,l} y_l \geq 0, \quad (17)$$

где $h_{n_r,l}$ — строки матрицы H , относящиеся к компонентам \bar{x} , на которые наложены ограничения. В отличие от неравенств $y_i \geq 0$, каждое из которых относится к единственной переменной, каждое неравенство (17) относится в общем случае ко всем переменным y_l . Соотношения (17), взятые для всех n_r , вместе с неравенствами $y_i \geq 0$, взятыми для всех l_r , представляют собой полную систему ограничений (9), наложенную на компоненты полного решения \bar{z} .

Граница допустимой области для x_{n_r} (то есть при $x_{n_r} = 0$) в общем виде описывается уравнением

$$\sum_{l=1}^L h_{n_r,l} y_l = f_{n_r}. \quad (18)$$

Если $L = 1$ (то есть существует единственная задаваемая координата y), то для каждого n_r соотношение (18) описывает точку. Это же соотношение описывает прямую линию на плоскости $\{y_1, y_2\}$ при $L = 2$, плоскость в трехмерном пространстве при $L = 3$ и гиперплоскость при еще большей размерности $\{y_i\}$.

Производные x_{n_r} по всем y_l равны

$$g_{n_r,l} = \frac{\partial x_{n_r}}{\partial y_l} = -h_{n_r,l}. \quad (19)$$

В простейшем случае $L = 1$ (y — одномерная величина), индекс l принимает единственное значение $l = 1$, и его можно опустить. Тогда каждое x_{n_r} равно нулю при $y = f_{n_r} / h_{n_r}$, а соот-

ветствующая величина g_{n_r} показывает направления возрастания всех x_{n_r} . Совмещение интервалов оси y , где все $x_{n_r} \geq 0$ (а также интервала $y \geq 0$, если такое ограничение имеется), дает значения y , при которых решение задачи существует. Не исключено, что это может быть единственное значение y , например, $y = 0$. Также не исключено, что интервал, на котором удовлетворяются все неравенства, отсутствует, и тогда решение полной задачи не существует.

Для $L > 1$ совокупность величин $\{h_{n_r,l}\} = \vec{h}_{n_r}$ есть вектор, перпендикулярный по отношению к n_r -й линейной границе (18), то есть ее нормаль, а $\vec{g}_{n_r} = \{g_{n_r,l}\}$ — градиент каждой величины x_{n_r} . В данной задаче градиент равен нормали к плоскости $x_{n_r} = 0$, взятой с обратным знаком: $\vec{g}_{n_r} = -\vec{h}_{n_r}$. Величина x_{n_r} равна нулю всюду на (18) и положительна в том полупространстве, в которое направлен градиент \vec{g}_{n_r} . С другой стороны, нормаль \vec{h}_{n_r} направлена в полупространство запрещенных значений $x_{n_r} < 0$.

Взаимное расположение границ (18) и сочетание этого расположения с направлениями градиентов (19) определяют существование и конфигурацию области допустимых решений системы (7) при $L \geq 2$.

Обозначим через N_r число ограничений (17) на компоненты x_{n_r} . Если $N_r \leq L$, то область в пространстве $\{y_i\}$, ограниченная линейными границами (18), будет бесконечной. При $N_r > L$ допустимая область может быть как бесконечной, так и конечной (то есть представлять собой многоугольник или многогранник) [Рокафеллар, 1973; Циглер, 2914]. Рассмотрим это на примерах небольшой размерности. На рисунках, представленных ниже, варианты расположения границ типа $x_{n_r} = 0$ помечены буквами. При обсуждении этих вариантов в тексте они обозначаются номером рисунка и соответствующей буквой (например, вариант 1в, 3а, 5д и т. д.). Для словосочетания «область допустимых значений \vec{y} » будем использовать также выражение «разрешенная область» или сокращение ОДЗ(y).

Пусть $L = 2$ — пространство $\{y_i\}$, представляет собой плоскость $\{y_1, y_2\}$. Если $N_r = 1$ (ограничение (17) наложено только на одну компоненту \vec{x}), то (18) — прямая, а допустимая область на $\{y_1, y_2\}$ — полуплоскость. При $N_r = 2$ возможны варианты, показанные на рисунке 1.

На всех рисунках номера n_r , соответствующие компонентам \vec{x} , на которые наложены ограничения, для простоты положены равными 1 и 2. Векторы \vec{g}_1 и \vec{g}_2 — градиенты x_1 и x_2 .

Вариант 1а изображает случай, когда линии ограничений пересекаются. Тогда, в зависимости от направлений градиентов, любой из углов на плоскости $\{y_1, y_2\}$ может быть разрешенной областью, но такая область для варианта 1а при любых \vec{g}_1 и \vec{g}_2 существует.

В случаях 1б–д прямые, соответствующие двум ограничениям, параллельны друг другу. Здесь направление градиентов играет большую роль.

В случае 1б ОДЗ(y) есть полуплоскость, как и при наличии только одного ограничения. Из двух ограничений должно быть оставлено одно, более сильное. В частности, может быть так, что обе границы совпадают, и тогда безразлично, какое из ограничений исключено как лишнее.

В случае 1в разрешенная область есть полоса между двумя границами, имеющая конечную ширину, но бесконечную длину.

В случае 1г ОДЗ(y) не существует, так как ограничения противоречат друг другу.

Наконец, вариант 1д представляет собой особый случай, так как обе границы совпадают, а градиенты направлены в противоположные стороны. Разрешенной областью является прямая, на которой обе подлежащие ограничению компоненты \vec{x} равны нулю (граничные значения допускаются неравенствами (9)). Это означает жесткую связь между компонентами \vec{y} .

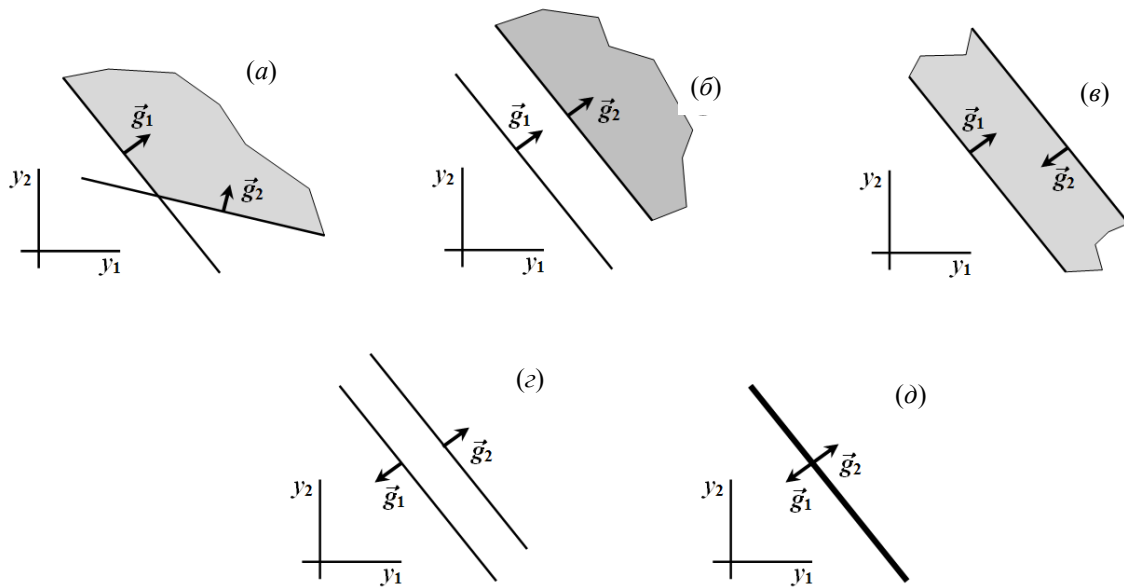


Рис. 1. Варианты области значений $\bar{y} = \{y_1, y_2\}$ при ограничениях на две компоненты \bar{x} (см. текст). Допустимая область затенена или изображена жирной линией

Возможные ограничения, наложенные непосредственно на y_1, y_2 или обе компоненты \bar{y} , на этом и последующих рисунках не показаны. Во всех случаях, когда разрешенная область \bar{y} существует, она имеет бесконечный размер по крайней мере в одном направлении. В случае 1d размерность решения на единицу меньше, чем размерность множества задаваемых переменных $\bar{y} = \{y_1, y_2\}$.

Варианты взаимного расположения ограничений (17) при $L = 2$ и $N_r = 3$, когда среди соответствующих прямых нет параллельных, показаны на рис. 2.

Существование и конфигурация ОДЗ(y) зависит от расположения прямых, описывающих ограничения $x_{n_r} \geq 0$, и от направлений градиентов x_{n_r} . Когда все три прямые не параллельны друг другу, они в общем случае ограничивают треугольник на плоскости $\{y_1, y_2\}$. Если все три градиента направлены внутрь треугольника, он является конечной областью допустимых значений \bar{y} (случай 2a). Если один или два градиента направлены вовне этого треугольника, то ОДЗ(y) существует, но лежит вне треугольника и является бесконечной. Нетрудно видеть, что при $L = 2$ и $N_r = 3$ возможны шесть таких вариантов: это внешние по отношению к треугольнику области между прямыми. На рис. 2 показаны два варианта из них — 2б и 2в. Если все три градиента направлены вовне треугольника, ограниченного прямыми $x_{n_r} = 0$ (вариант 2з), то ОДЗ(y) не существует.

Если все три прямые пересекаются в одной точке, то возможны два варианта. В случае 2д разрешенная область содержит только эту точку (то есть она нульмерна). В случае 2е, напротив, существует двумерная ОДЗ(y) (угол на плоскости $\{y_1, y_2\}$). Для ее существования необходимо, чтобы границы одновременно удовлетворяли двум требованиям: 1) градиенты двух из x_{n_r} (на рис. 2 это \bar{g}_3 и \bar{g}_2) направлены внутрь одного из углов, образованных соответствующими прямыми; 2) третья прямая проходит внутри этого же угла.

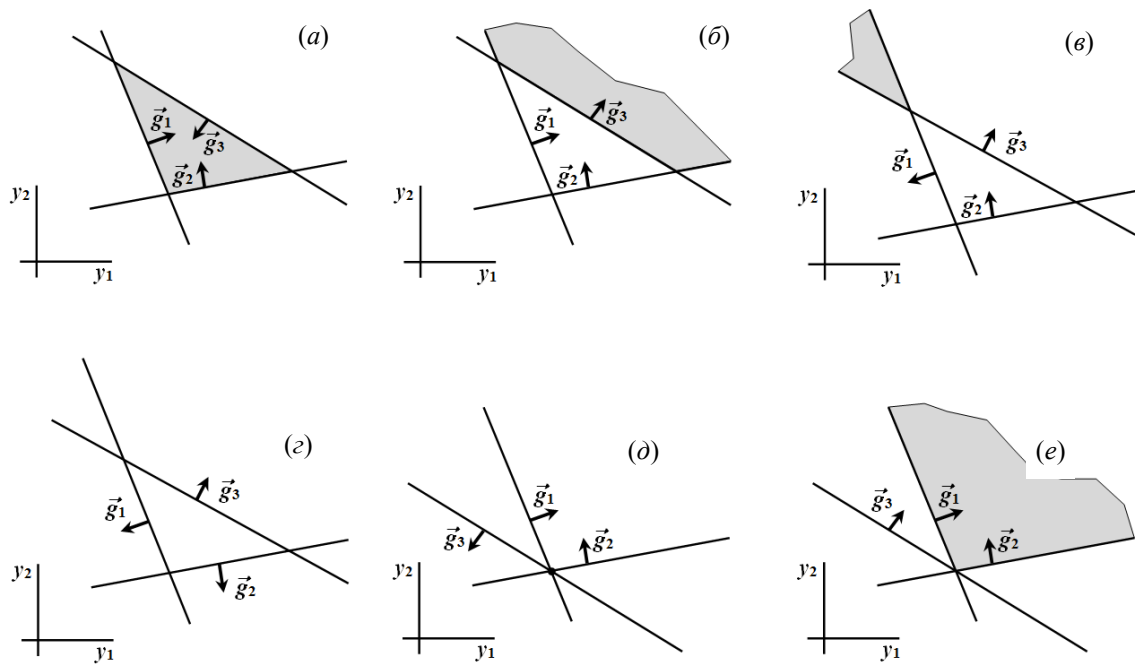


Рис. 2. Варианты области значений $\vec{y} = \{y_1, y_2\}$ при ограничениях на три компоненты \vec{x} , когда среди соответствующих прямых нет параллельных (см. текст). Допустимая область затенена или изображена жирной точкой

На рисунке 3 показаны варианты, когда $L = 2$ и $N_r = 3$, и при этом две прямые из трех параллельны друг другу.

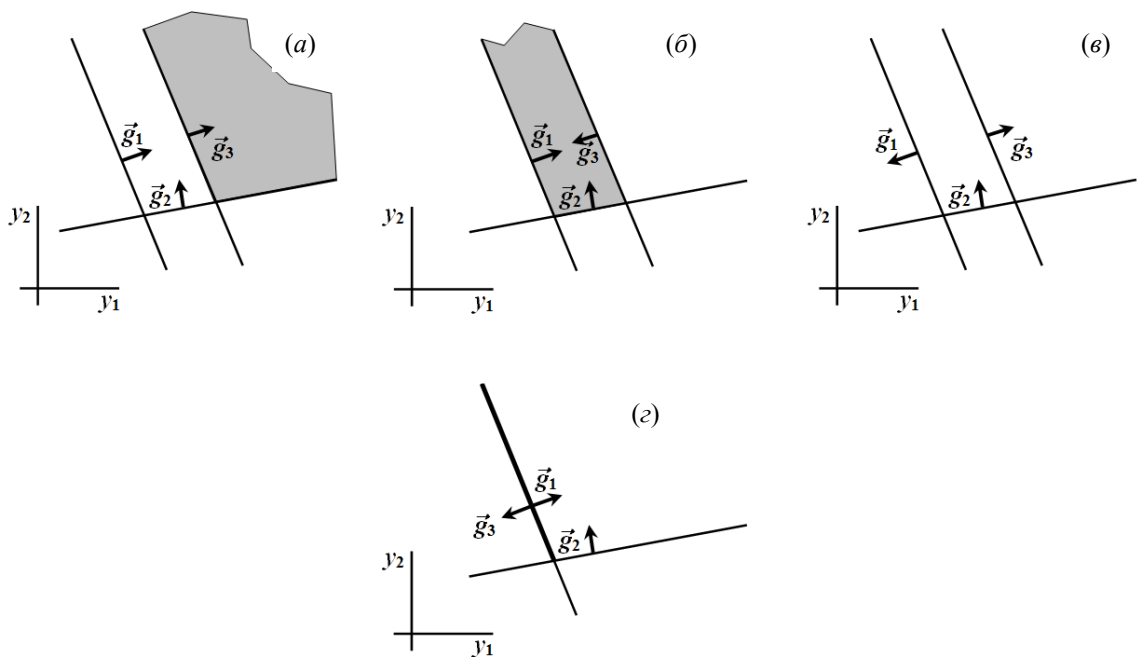


Рис. 3. Варианты области значений $\vec{y} = \{y_1, y_2\}$ при ограничениях на три компоненты \vec{x} , когда две из трех прямых параллельны друг другу (см. текст). Допустимая область затенена или изображена жирной линией

На рисунках 3а–г влияние ограничений, изображаемых двумя параллельными прямыми, такое же, как на рисунке 1б–д, где третьего ограничения нет. Здесь в случае 3в, как и случае 1г, ОДЗ(y) не существует, независимо от расположения третьей прямой. Для остальных вариантов решение существует, но третье ограничение запрещает часть области, разрешаемой первыми двумя.

На рисунке 3 не показан случай, когда все три прямые на плоскости $\{y_1, y_2\}$, соответствующие ограничениям $x_n \geq 0$, параллельны. Тогда, по аналогии с рисунком 1, возможны ситуации, когда разрешенная область — полуплоскость, полоса, прямая либо не существует. Полуплоскость имеет место, когда все три градиента «смотрят» в одну сторону. Полоса — когда две прямые и два градиента расположены как в случае 1в, третья прямая находится вне разрешенной полосы, а третий градиент направлен в сторону этой полосы. Допустимая область в виде прямой существует тогда, когда две прямые и соответствующие градиенты расположены, как на рисунке 1д, а третий градиент направлен в сторону этой прямой. Если же третий градиент направлен в сторону, противоположную от полуплоскости, полосы или прямой, описываемых двумя другими ограничениями, то ОДЗ(y) не существует.

Варианты взаимного расположения ограничений (17) при $L = 2$ и $N_r = 4$ показаны на рисунках 4 и 5.

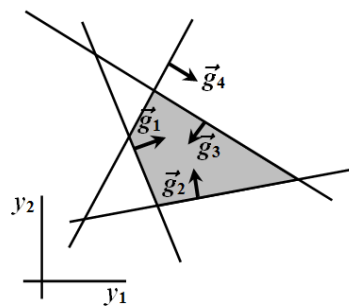


Рис. 4. Расположение ограничений на переменные y_1, y_2 , накладываемых условиями $x_n \geq 0$ в случае, когда число этих условий равно 4, а среди прямых–границ разрешенных областей нет параллельных друг другу (см. текст). Показан один из вариантов направлений градиентов величин x_n , затенен соответствующий вариант допустимой области y_1, y_2

Из рисунка 4 видно, что в случае, когда все прямые не параллельны друг другу, они ограничивают несколько областей, среди которых есть треугольные и может существовать один четырехугольник. В зависимости от направлений градиентов разрешенной может быть любая конечная или бесконечная область, ограниченная двумя, тремя или (если существует четырехугольник) четырьмя лучами или отрезками. Здесь уже достаточно много вариантов.

Из рисунка 5 видно, что варианты 5а–г аналогичны вариантам 3а–г (рис. 3), где изображена ситуация при наличии трех ограничений. Четвертая прямая (соответствующий градиент — \vec{g}_2) изображает еще одно ограничение на допустимую область переменных y_1, y_2 , отсутствующее при $N_r = 3$ (рис. 3). Вариант 5д изображает ситуацию, когда четвертое ограничение (соответствующий градиент — \vec{g}_4) никак не изменяет ситуацию по сравнению с той, когда его нет. Напротив, вариант 5е коренным образом все изменяет — введение четвертого ограничения полностью уничтожает какие бы то ни было решения задачи, возможные в его отсутствие.

Можно видеть, что рисунок 5 не исчерпывает возможные варианты формы и расположения допустимой области значений задаваемых переменных y_1, y_2 . Например, если обратить на-

правление градиента \vec{g}_2 , то в варианте 5г разрешенной областью станет другая часть той же прямой. В варианте 5д это тогда будет полоса между двумя параллельными прямыми, ограниченная еще и с верхнего торца; граница с градиентом \vec{g}_4 станет играть важную роль, а при другом направлении \vec{g}_2 она не влияла вообще. В варианте 5е появится допустимая область, которой не было при направлениях градиентов, указанных на рисунке 5.

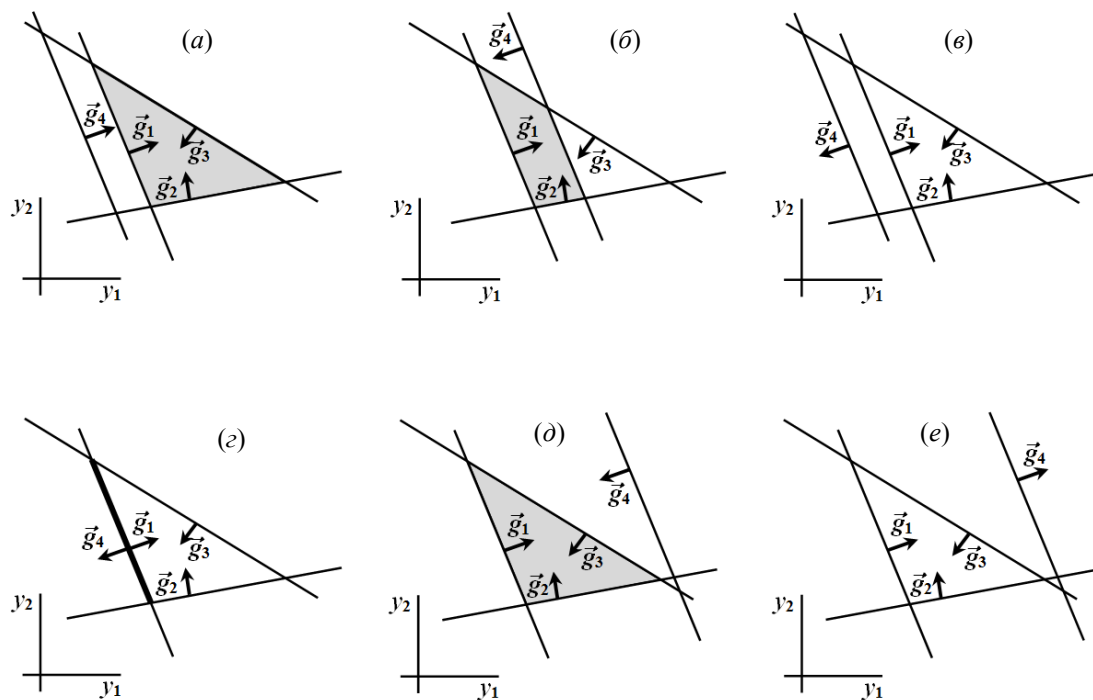


Рис. 5. Расположение ограничений на переменные y_1, y_2 , накладываемых условиями $x_n \geq 0$ в случае, когда число этих условий равно 4, а среди прямых-границ разрешенных областей есть две параллельные друг другу (см. текст). Допустимая область затенена или изображена жирной линией

Если L — размерность задаваемого вектора \vec{y} — равна 3, то каждая граница допустимой области \vec{y} , то есть, $x_n = 0$, будет уже не прямой, а плоскостью. При еще больших значениях L границы станут гиперплоскостями. Напомним, что к ограничениям на \vec{y} , вытекающим из условий $x_n \geq 0$, следует присоединить прямые ограничения на часть компонент \vec{y} , а именно $y_l \geq 0$ (см. (9)). Все это усложняет задачу поиска допустимых значений \vec{y} . Возникает вопрос о рациональном способе осуществления такого поиска.

Поиск допустимых значений задаваемых переменных

Решение основной задачи дается выражением (16). Базисные векторы $\vec{y}^{(\alpha)}$ задаются произвольно, но в пределах, отвечающих неравенствам (9) и физическим ограничениям (например, значения задаваемых входных и выходных метаболических потоков должны соответствовать возможностям ферментных систем клеток). После задания вектора \vec{y} вектор \vec{x} находится из системы уравнений (11) (ее детальный вид — (12)). Затем находятся компоненты b_j для присоединенной задачи, если она содержится в общей задаче.

Главной проблемой при поиске решения является нахождение области допустимых значений \vec{y} (далее будем обозначать ее сокращенно как ОДЗ(y)). Поиск этой области требует зада-

ния начальной точки для вычислений. Она может находиться внутри ОДЗ(y), на ее границе или вне этой области, но вблизи от нее. В последних двух случаях найти точку, лежащую внутри ОДЗ(y), можно, основываясь, например, на направлениях градиентов \vec{g}_{n_r} и используя точку \vec{y} , полученную на первом шаге вычислений, как отправную. Сканирование пространства $\{y\}$ при этом будет сделано с меньшими вычислительными затратами, чем из некоторой произвольно назначенной стартовой точки.

Если точка с координатами \vec{y}_1 , находящаяся внутри ОДЗ(y), найдена, то линейно независимые друг от друга векторы $\vec{y}^{(\alpha)}$ могут быть образованы вариациями компонент \vec{y}_1 вдоль осей y_i : $\vec{y}^{(\alpha)} = \vec{y}_1 + \Delta\vec{y}^{(\alpha)}$, где каждое отклонение $\Delta\vec{y}^{(\alpha)}$ делается только вдоль одной из координат, причем величина каждого отклонения конечна и не мала. Полное решение задачи требует нахождения всех частей границы допустимой области значений \vec{y} . Но для качественного исследования (например, определения базисных метаболических путей) может быть достаточно нахождение набора точек $\vec{y}^{(\alpha)}$.

Предлагаемый здесь подход основан на использовании свойств совокупности ограничений на ОДЗ(y), описанных в предыдущем разделе. Подход работает в пространстве любой размерности, то есть независимо от того, являются ли границы области $x_{n_r} \geq 0$ прямыми линиями на плоскости, плоскостями в трехмерном пространстве или гиперплоскостями в пространстве большей размерности, чем 3. Ниже будем условно называть такие геометрические объекты общим термином «плоскость».

Исключение лишних ограничений

Прежде всего, следует выяснить, существуют ли в данной задаче параллельные друг другу границы. Это нетрудно сделать, сравнивая направления \vec{g}_{n_r} — градиентов x_{n_r} . Если угол между двумя \vec{g}_{n_r} равен нулю, они равнонаправленные; если он равен π , градиенты направлены навстречу друг другу, но вдоль одной прямой. Далее, если несколько \vec{g}_{n_r} «смотрят» в одну сторону, то нужно выбрать границу, соответствующую наиболее сильному ограничению (см., например, рис. 1б), а остальные ограничения в дальнейшем анализе не рассматривать. Тогда вместо каждой группы параллельных границ с равнонаправленными градиентами величин x_{n_r} останется одно ограничение.

Далее следует выяснить, нет ли границы, параллельной данной границе, но с противоположно направленным вектором \vec{g}_{n_r} . (Вторая граница также должны быть получена исключением лишних границ с её направлением градиента.) Если таковая имеется, то возникает вопрос о взаимном расположении этих двух границ. Возможны три ситуации.

1. Если ограничения противоречат друг другу (например, как на рис. 1з), то ОДЗ(y) не существует, каковы бы ни были прочие ограничения, и тогда решение общей задачи также не существует.

2. Если эти границы отодвинуты друг от друга так, что соответствующие градиенты \vec{g}_{n_r} направлены внутрь области между ними, то эта область, включая сами границы, является допустимой, то есть, ОДЗ(y). В частном случае, когда $L = 2$, это случай рисунка 1в.

3. Промежуточный вариант, когда две такие границы совпадают (например, случай рис. 1д). Тогда только прямая, плоскость или гиперплоскость (в зависимости от размерности пространства $\{y\}$), описывающая две совпадающие границы, является допустимой областью для \vec{y} . В таком случае ОДЗ(y) описывается не неравенством $x_{n_r} \geq 0$, а равенством $x_{n_r} = 0$. Это жесткая связь между компонентами задаваемого вектора \vec{y} .

Если в данной задаче существует хотя бы одна группа параллельных между собой границ, вышеизложенный анализ позволит заметно упростить дальнейший поиск области допустимых значений \bar{y} . Упрощение будет тем более существенным, чем больше параллельных границ в одной группе и чем больше таких групп. Из сказанного следует, что ОДЗ(y) не может состоять из двух или более отделенных друг от друга параллельных областей — полос или пространственных слоев.

В зависимости от задачи жесткие связи могут не существовать вообще или может оказаться, что связь не одна, а их несколько. При этом плоскости, описывающие такие жесткие связи, обязаны пересекаться, иначе не будет существовать значений \bar{y} , удовлетворяющих одновременно нескольким жестким связям. Тогда ОДЗ(y) может находиться только в области пересечения таких плоскостей. Прочие ограничения, не совпадающие с жесткими связями и поэтому имеющие вид неравенств, могут привести к тому, что ОДЗ(y) составляет лишь часть области пересечения плоскостей, соответствующих жестким связям. В любом случае при наличии хотя бы одной жесткой связи размерность области в пространстве $\{y\}$, в которой нужно искать ОДЗ(y), снижается.

Далее номера n_r остающихся ограничений разобьем на две группы: n_{r_0} для жестких связей между компонентами вектора \bar{y} и n_{r_1} для неравенств вида (17). Числа номеров n_{r_0} и n_{r_1} обозначим как N_{r_0} и N_{r_1} . Все $x_{n_{r_0}}$ при \bar{y} , принадлежащих пересечению жестких связей, имеют фиксированные значения, которые в том числе могут быть нулевыми.

Прямой поиск точки, находящейся в области допустимых значений \bar{y}

Вспользуемся параметрическим уравнением для перпендикуляра, опущенного на плоскость, описывающую ограничение (17) на $x_{n_{r_1}}$ (см. приложение 1). Задача ставится так: нужно найти координаты точки \bar{y}_1 такой, что для всех перпендикуляров, опущенных из этой точки на все плоскости, параметр α будет иметь одно и то же значение. Для любой из этих плоскостей уравнение (П.4) может быть переписано как $\vec{h}_{n_r} \bar{y}_1 + \alpha |\vec{h}_{n_r}|^2 = f_{n_r}$ (в качестве \bar{y}^{sp} здесь выступает точка \bar{y}_1).

Соответствующая система уравнений имеет вид

$$\sum_{l=1}^L h_{n_r,l} y_{1l} + \left[\sum_{l=1}^L (h_{n_r,l})^2 \right] \alpha = f_{n_r} \quad \text{для всех } n_{r_1}. \quad (20)$$

Число неизвестных здесь равно $L+1$ (все компоненты \bar{y}_1 плюс неизвестная α). Число уравнений (20) равно N_{r_1} . Сюда следует добавить N_{r_0} уравнений для жестких связей между переменными \bar{y} , которые действуют и в отношении точки \bar{y}_1 (в эти уравнения параметр α не входит):

$$\vec{h}_{n_{r_0}} \bar{y}_1 = \sum_{l=1}^L h_{n_{r_0},l} y_{1l} = f_{n_{r_0}} \quad \text{для всех } n_{r_0}. \quad (21)$$

Итак, общее число уравнений для вектора $\begin{pmatrix} \bar{y}_1 \\ \alpha \end{pmatrix}$ равно $N_r = N_{r_0} + N_{r_1}$. Решение существует, если число уравнений системы (20)–(21) не больше, чем $L+1$. При этом если существует пересечение всех плоскостей, как с номерами n_{r_0} , так и n_{r_1} , то $\alpha = 0$ (то есть расстояние от \bar{y}_1 до всех ограничительных плоскостей равно нулю и точка \bar{y}_1 принадлежит упомянутому пересечению). Область положительных $x_{n_{r_1}}$ указывается направлениями $\vec{g}_{n_{r_1}}$, а вернее проекциями этих градиентов на область пересечения жестких связей (см. ниже).

Поиск методом наименьших квадратов

Этот подход может дать результат, лежащий как внутри ОДЗ(y) (то есть где $x_{n_i} > 0$), так и вне ОДЗ(y), но вблизи этой области.

Рассмотрим случай конечной области значений \bar{y} , ограниченной со всех сторон оставленными границами. Выдвинем требование минимальности суммы квадратов расстояний искомой точки в пространстве $\{y\}$ от всех границ $x_{n_i} \geq 0$. Такой подход аналогичен поиску параметров линейной регрессии методом наименьших квадратов [Брандт, 2003]. Примеры, изображенные на рисунках 2, 4 и 5, показывают, что найденная таким способом точка будет лежать или в искомой допустимой области, или рядом с ней. Например, как видно на рисунке 4, треугольник левее и выше ОДЗ(y), очевидно, обеспечит более низкое значение суммы квадратов отклонений от трех ограничительных линий, чем затененная область, изображающая ОДЗ(y). Поэтому обозначим точку, искомую таким методом, как \bar{y}_0 , в отличие от \bar{y}_1 .

Начнем с ситуации, когда жесткие связи между компонентами \bar{y} отсутствуют. Выражая расстояние от \bar{y}_0 до каждой из границ в виде (П.5) (см. приложение 1), получаем, что эта точка должна быть найдена из условий минимума величины

$$\Pi^2 = \sum_{n_r} P_{n_r}^2 = \sum_{n_r} (f_{n_r} - \bar{h}_{n_r} \bar{y}_0)^2 / |\bar{h}_{n_r}|^2, \tag{22}$$

где $\bar{h}_{n_r} \bar{y}_0 = \sum_{l=1}^L h_{n_r,l} y_{0l}$, $|\bar{h}_{n_r}|^2 = \sum_{l=1}^L (h_{n_r,l})^2$.

Уравнения для нахождения точки \bar{y}_0 , удовлетворяющей вышеупомянутому требованию, следуют из условий

$$\frac{\partial \Pi^2}{\partial y_{0m}} = 0, \quad m = 1, 2, \dots, L. \tag{23}$$

Подставляя сюда (22), получаем систему из L линейных уравнений для L неизвестных координат искомой точки \bar{y}_0 :

$$\sum_{l=1}^L M_{ml} y_{0l} = B_m, \quad m = 1, 2, \dots, L, \tag{24}$$

$$\text{где } M_{ml} = \sum_{n_r} \frac{h_{n_r,m} h_{n_r,l}}{(\bar{h}_{n_r})^2}, \quad B_m = \sum_{n_r} \frac{f_{n_r} h_{n_r,m}}{(\bar{h}_{n_r})^2}. \tag{25}$$

Если $\text{rank}(M_{ml}) = L$, то \bar{y}_0 находится однозначно. Ситуация, когда $\text{rank}(M_{ml}) < L$, обсуждается ниже.

В (22)–(24) не учтены прямые ограничения (9) на некоторые компоненты \bar{y} . Проверка соблюдения соотношений $y_{l_i} \geq 0$ может быть сделана после того, как точка \bar{y}_0 , удовлетворяющая требованиям $x_{n_i} \geq 0$, найдена. С другой стороны, если считается целесообразным включить границы $y_{l_i} = 0$ в число объектов, «притягивающих» точку \bar{y}_0 , то минимизируемая функция должна иметь вид

$$\Pi^2 = \sum_{n_r} P_{n_r}^2 + \sum_{l_i} y_{0l_i}^2. \tag{26}$$

Уравнение (24) при этом остается справедливым, как и B_m из (25). Но M_{ml} принимает следующий вид:

$$M_{ml} = \sum_{n_r} \left(\frac{h_{n,m} h_{n,l}}{(\bar{h}_{n_r})^2} + \delta_{ll'} \delta_{lm} \right), \quad (27)$$

где δ — символ Кронекера, равный 1, если значения обоих индексов равны друг другу, и 0 — в остальных случаях.

Если имеются жесткие связи между компонентами \bar{y} , формулировка задачи должна быть еще раз изменена, поскольку искать в этом случае нужно условный минимум. Жесткие связи описываются соотношениями (21), которые здесь будем использовать в форме $f_{n_{r_0}} - \sum_{l=1}^L h_{n_{r_0}l} y_{0l} = 0$. Им должны удовлетворять в том числе и значения компонент искомого вектора \bar{y}_0 . Используем для решения метод неопределенных множителей Лагранжа. Согласно этому методу левые части уравнений жестких связей добавляются к функции, условный экстремум которой ищется с неопределенными множителями. Последние сами по себе не имеют значения для основной задачи, но они вместе с координатами экстремума \bar{y}_0 входят как компоненты в вычисляемый вектор. В нашем случае минимизируемая функция имеет вид

$$\Pi^2 = \sum_{n_{r_1}} P_{n_{r_1}}^2 + \sum_{l_r} y_{0l_r}^2 + \sum_{n_{r_0}} \lambda_{n_{r_0}} \left(f_{n_{r_0}} - \sum_{l=1}^L h_{n_{r_0}l} y_{0l} \right). \quad (28)$$

Уравнения (23) для минимума Π^2 по y_{0m} принимают вид

$$\sum_{l=1}^L M_{ml} y_{0l} - \frac{1}{2} \sum_{n_{r_0}m} h_{n_{r_0}m} \lambda_{n_{r_0}} = B_m, \quad m = 1, 2, \dots, L, \quad (29)$$

где матрица M_{ml} имеет вид (27), правые части B_m такие же, как в (25). К системе из L линейных уравнений для y_{0l} и $\lambda_{n_{r_0}}$ добавляются N_{r_0} уравнений (21), также линейных. Вектор неизвестных в такой постановке задачи есть $\begin{pmatrix} \bar{y}_0 \\ \bar{\lambda} \end{pmatrix}$, а матрица системы уравнений — конкатенация

двух матриц в (29): $\left(M_{ml}, -\frac{1}{2} h_{n_{r_0}m} \right)$. Общее число как уравнений, так и неизвестных равно $L + N_{r_0}$. Поэтому общая система уравнений (29) и (21) имеет единственное решение, если ранг этой системы равен $L + N_{r_0}$.

Решение будет единственным, когда область, которой принадлежит точка \bar{y}_0 , является конечной, то есть когда граничные плоскости образуют замкнутую поверхность. Рисунки 2б и 2в, например, показывают, что такая область не обязательно является ОДЗ(y). Как уже говорилось, если точка \bar{y}_0 не принадлежит ОДЗ(y), с нее удобно начинать дальнейший поиск точки \bar{y}_1 , которая действительно принадлежит ОДЗ(y). Кроме пошагового поиска, возможны и более быстрые методы. Например, если, как в случае 2б, для точки \bar{y}_0 , находящейся внутри треугольника, нарушается единственное ограничение, то точку \bar{y}_1 можно получить, зеркально отразив точку \bar{y}_0 относительно нарушенной границы.

Наконец, если область, искомая методом наименьших квадратов, неограниченно простирается в пространстве $\{y\}$ в одном или нескольких направлениях, то ранг системы уравнений

для $\begin{pmatrix} \bar{y}_0 \\ \bar{\lambda} \end{pmatrix}$ будет меньше, чем $L + N_{r_0}$. Тогда можно исключить эту трудность, зафиксировав значения соответствующего (см. ниже) числа переменных y_l , то есть наложив на них жесткую связь простейшего вида. В частности, может оказаться удобным положить их равными нулю. Возможны также и содержательные связи, имеющие «физический» смысл. Число таких условий должно быть равно разности между размером и рангом системы уравнений для $\begin{pmatrix} \bar{y}_0 \\ \bar{\lambda} \end{pmatrix}$.

Направления наискорейшего возрастания величин x_{n_1} в подпространстве, заданном жесткими связями между y_l

Направление наискорейшего возрастания каждой из величин x_{n_1} в полном пространстве $\{y\}$ задается градиентом $\bar{g}_{n_1} = \nabla_{\bar{y}} x_{n_1} = -\bar{h}_{n_1}$ (см. (19)). Существование одной или нескольких плоскостей, описывающих жесткие связи между переменными y_l , задает подпространство, соответствующее одной плоскости (если жесткая связь одна) или пересечению таких плоскостей (если их несколько). Система уравнений для него имеет вид $\bar{h}_{n_0} \bar{y} = f_{n_0}$ (см. (21)). Все изменения x_{n_1} , в частности при поиске ОДЗ(y), должны производиться в этом подпространстве. Направление и величина наискорейшего увеличения x_{n_1} («глядя» из полного пространства $\{y\}$) тогда задаются вектором $\bar{g}_{n_1}^{pr}$ — проекцией \bar{g}_{n_1} на пересечение плоскостей жестких связей. Из формул (П.10) и (П.9) (приложение 2) следует, что

$$\bar{g}_{n_1}^{pr} = \bar{g}_{n_1} + \sum_{n_{r_0}} \lambda_{n_{r_0}}^{(g)} \bar{h}_{n_{r_0}}, \quad (30)$$

где $\lambda_{n_{r_0}}^{(g)}$ находятся из системы линейных уравнений

$$\sum_{n_{r_0}} \lambda_{n_{r_0}}^{(g)} \bar{h}_{m_{r_0}} \bar{h}_{n_{r_0}} = -\bar{h}_{m_{r_0}} \bar{g}_{n_1}. \quad (31)$$

Следует обратить внимание на то, что в (30) и (31) присутствуют две нумерации компонент векторов. Номера n_{r_0} и m_{r_0} относятся к нормальным плоскостям, описывающим подпространство, в котором соблюдаются сразу все жесткие связи между y_l , и к связанным с ними коэффициентам λ . Номера n_{r_1} относятся к проецируемым градиентам \bar{g}_{n_1} и к полученным их проекциям $\bar{g}_{n_1}^{pr}$. Коэффициенты $\lambda_{n_{r_0}}^{(g)}$ зависят как от плоскостей, на пересечение которых производится проецирование (номера n_{r_0} и m_{r_0} в (31)), так и от проецируемого вектора \bar{g}_{n_1} (значок (g) у λ). Детали операции такого проецирования изложены в приложении 2.

Полученные векторы $\bar{g}_{n_1}^{pr}$ могут быть использованы для движения в полном пространстве $\{y\}$ в сторону области допустимых значений \bar{y} (то есть тех, которые обеспечивают соблюдение требования неотрицательности компонент \bar{x}_{n_1}), не выходя при этом из подпространства, в котором соблюдаются все жесткие связи вида (21).

Базисные решения биохимического примера

В таблицах 4 и 5 приведены решения для конкретной задачи, сформулированной в разделе «Пример, поясняющий остановку задачи». Жирной рамкой обведены столбцы обеих таблиц, описывающие базисные решения — потоки через реакции $\vec{z}^{(\alpha)}$ (таблица 4) и потоки обмена рассматриваемой метаболической системы с ее окружением \vec{b} . Напомним, что скорости реакций \vec{z} являются решениями основной задачи, в которой число линейно независимых уравнений I (в данном примере $I = 18$) меньше K — числа неизвестных, то есть компонент вектора \vec{z} (здесь $K = 20$). Число линейно независимых решений $\vec{z}^{(\alpha)}$ равно $K - I + 1 = 3$. Соответственно, мы имеем в данном случае три базисных столбца в таблицах 4 и 5. Значения компонент вектора \vec{b} , соответствующие основной задаче, фиксированы ее формулировкой. В то же время компоненты \vec{b} , соответствующие присоединенной задаче, зависят от \vec{z} и поэтому могут различаться для разных $\vec{z}^{(\alpha)}$. Следовательно, полный вектор \vec{b} для каждого α -го варианта решения нужно обозначить тем же индексом: $\vec{b}^{(\alpha)}$.

Столбец, не обведенный жирной линией в таблицах 4 и 5, является примером линейной комбинацией базисных решений: $\vec{z} = -0.5\vec{z}^{(1)} + 0.5\vec{z}^{(2)} + \vec{z}^{(3)}$ и $\vec{b} = -0.5\vec{b}^{(1)} + 0.5\vec{b}^{(2)} + \vec{b}^{(3)}$.

Таблица 4. Решения для потоков \vec{z} через реакции рассматриваемой ферментной системы. Базисные решения $\vec{z}^{(\alpha)}$ показаны в столбцах внутри жирной рамки. Четвертый столбец является одной из их линейных комбинаций: $\vec{z} = -0.5\vec{z}^{(1)} + 0.5\vec{z}^{(2)} + \vec{z}^{(3)}$ (подробности в тексте)

k	Реакции	α			Линейная комбинация
		1	2	3	
1	Электрон-транспортная цепь (комплекс I)	3	3	3	3
2	Электрон-транспортная цепь (остальная часть)	3	3	3	3
3	Пируватдекарбоксилаза	0	2	0	1
4	Пируватдегидрогеназный комплекс	2	0	1	0
5	Пируваткарбоксилаза	0	0	1	1
6	Цитратсинтаза	1	1	1	1
7	Аконитаза (стадия 1)	1	1	1	1
8	Аконитаза (стадия 2)	1	1	1	1
9	Изоцитратдегидрогеназа (стадия 1)	0	0	1	1
10	Изоцитратдегидрогеназа (стадия 2)	0	0	1	1
11	α -кетоглутаратдегидрогеназа	0	0	1	1
12	Сукцинил-КоА-синтаза (ГТФ)	0	0	-1	-1
13	Сукцинатдегидрогеназа	0	0	0	0
14	Фумараза	0	0	0	0
15	Малатдегидрогеназа	1	1	0	0
16	Ацетальдегиддегидрогеназа	0	2	0	1
17	Изоцитратлиаза	1	1	0	0
18	Малатсинтаза	1	1	0	0
19	Алкогольдегидрогеназа	0	0	0	0
20	Лактатдегидрогеназа	0	0	0	0

Отметим, что все потоки через необратимые реакции (индекс необратимости 1, см. таблицу 1) в базисных решениях неотрицательны. Линейная комбинация в приведенном примере также отвечает этому требованию. В общем случае при задании набора коэффициентов $X^{(\alpha)}$ соблюдение ограничений (9) для каждой комбинации базисных решений должно проверяться.

Таблица 5. Величины потоков \bar{b} , осуществляющих обмен рассматриваемой ферментной системы с ее окружением. Значения компонент \bar{b} даны для общей задачи, состоящей из основной (эти компоненты \bar{b} показаны под двойной разделительной линией) и присоединенной (над этой линией). Столбцы таблицы 5 соответствуют столбцам таблицы 4 (подробности в тексте)

j ↓	i ↓	Метаболиты	α			Линейная комбинация	
			1	2	3		
1		O ₂	-1.5	-1.5	-1.5	-1.5	Вычисленные потоки (присоединенная задача)
2		CO ₂	2	2	2	2	
3		H ₂ O	8.5	8.5	8.5	8.5	
4		H ⁺	0	0	0	0	
5		H ₃ PO ₄	-7.5	-7.5	-7.5	-7.5	
6		АТФ	7.5	7.5	6.5	6.5	
7		АДФ	-7.5	-7.5	-6.5	-6.5	
8		ГТФ	0	0	1	1	
9		ГДФ	0	0	-1	-1	
10	1	НАДН	0	0	0	0	Заданные потоки (основная задача)
11	2	НАД ⁺	0	0	0	0	
12	3	убихинол	0	0	0	0	
13	4	убихинон	0	0	0	0	
14	5	КоА	0	0	0	0	
15	6	ацетил-КоА	0	0	0	0	
16	7	сукцинил-КоА	0	0	0	0	
17	8	пируват	-2	-2	-2	-2	
18	9	ацетальдегид	0	0	0	0	
19	10	оксалоацетат	0	0	0	0	
20	11	цитрат	0	0	0	0	
21	12	<i>цис</i> -аконитат	0	0	0	0	
22	13	изоцитрат	0	0	0	0	
23	14	оксалосукцинат	0	0	0	0	
24	15	α -кетоглутарат	0	0	0	0	
25	16	сукцинат	1	1	1	1	
26	17	фумарат	0	0	0	0	
27	18	L-малат	0	0	0	0	
28	19	глиоксилат	0	0	0	0	
29	20	этанол	0	0	0	0	
30	21	лактат	0	0	0	0	

На рисунках 6 и 7 показаны метаболические пути для базисных решений $\bar{z}^{(1)}$ и $\bar{z}^{(3)}$ соответственно. Верхние части рисунков изображают низкомолекулярные части метаболических путей, нижние — электрон-транспортную цепь. Потоки через необратимые реакции изображены двойными линиями, через обратимые — обычными линиями. Цифры показывают значения потоков (в относительных единицах). Отличие этих двух базисных путей состоит в том, что $\bar{z}^{(1)}$ включает так называемый глиоксилатный шунт цикла Кребса, а $\bar{z}^{(3)}$ вместо изоцитратлиазы и малатсинтазы использует дальнейшие реакции этого цикла.

Метаболические пути для базисного вектора $\bar{z}^{(2)}$ можно легко представить себе с помощью рисунка 6, если на нем заменить пируватдегидрогеназный комплекс (ПДГК) на две отдельные реакции, катализируемые пируватдекарбоксилазой (ПДК) и ацетальдегиддегидрогеназой (ААДГ). В целом эти две реакции выполняют тот же процесс, что и ПДГК, но эти ферменты различаются: согласно международной классификации ферментов ПДГК имеет номер ЕС 1.2.4.1, ПДК — ЕС 4.1.1.1, ААДГ — ЕС 1.2.1.10. На рисунке 8 изображены оба варианта

ферментативных реакций, осуществляющих преобразование



Вариант $\bar{z}^{(2)}$ получается из рис. 6 заменой верхнего фрагмента рисунке 8 на нижний.

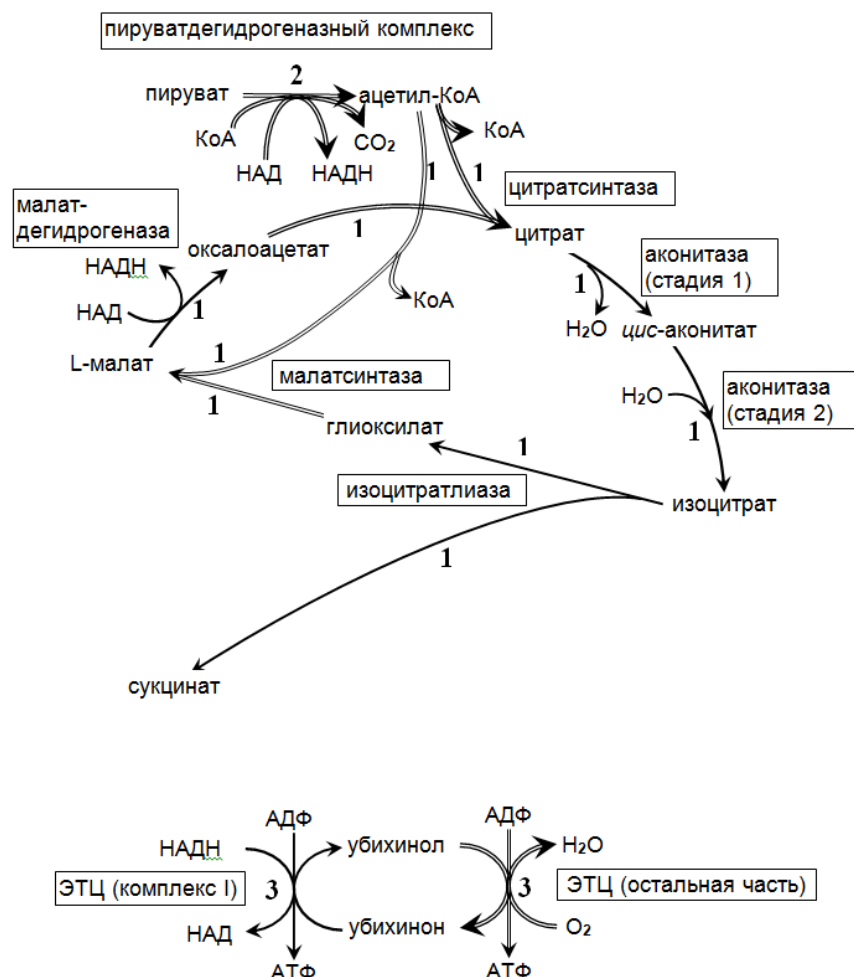


Рис. 6. Метаболические пути, соответствующие базисному вектору $\bar{z}^{(1)}$ (подробности в тексте)

Отметим, что интуитивно представляется, что в этой задаче должны быть четыре базисных решения, так как совокупность реакций можно разделить на две последовательные части (пируват \rightarrow ацетил-КоА и всё остальное), каждая из которых имеет два варианта. Однако строгий анализ показал, что линейно независимых решений только три, а остальные могут быть получены как их суммы с весовыми коэффициентами. С содержательной же точки зрения (в данном случае с точки зрения биохимика) число решений, имеющих самостоятельное «физическое» значение, может быть больше, чем число базисных решений.

Обсуждение

Описанные выше закономерности важны для разработки алгоритмов поиска решений неполных систем линейных уравнений с ограничениями на допустимые значения переменных. Именно эти ограничения создают наибольшие трудности нахождения решений. В статьях

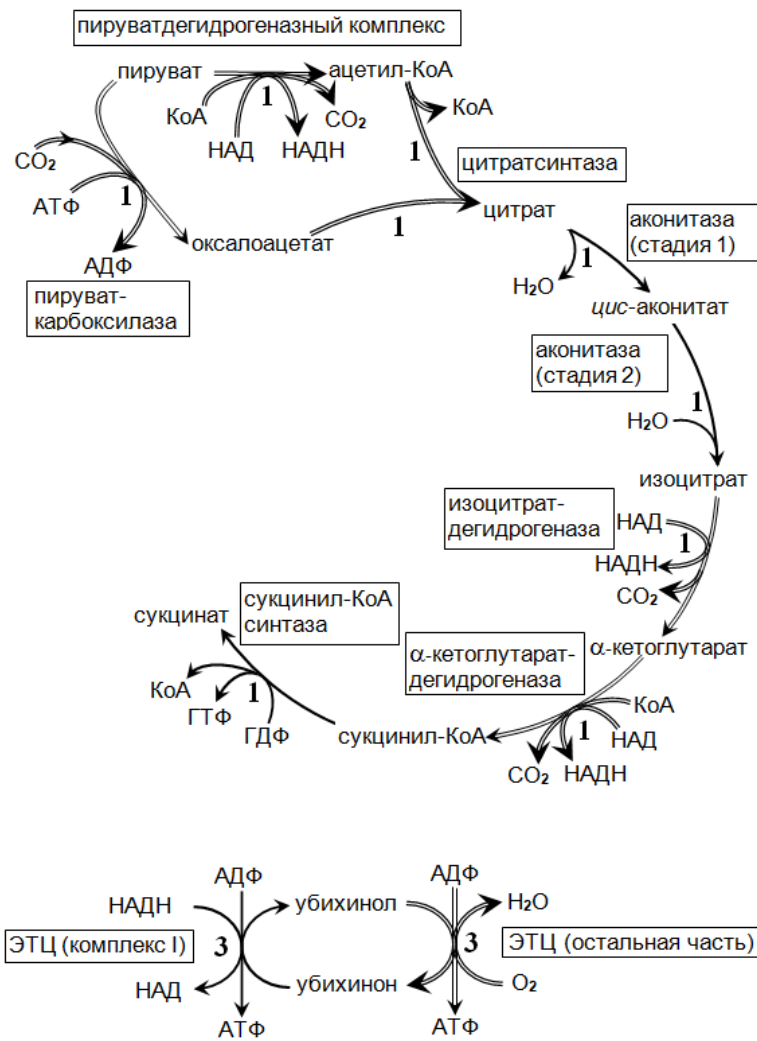


Рис. 7. Метаболические пути, соответствующие базисному вектору $\bar{z}^{(3)}$ (подробности в тексте)

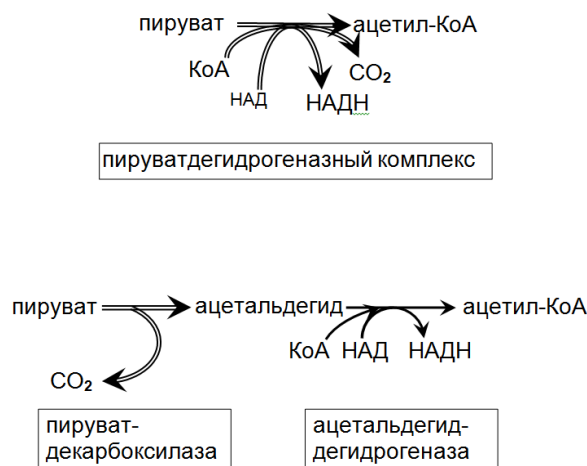


Рис. 8. Два варианта реакций, осуществляющих процесс, согласно уравнению (32) (подробности в тексте)

[Schilling et al., 1999; Papin J. A., 2004] приведен обзор существующих методов решения подобных задач в области стехиометрии метаболизма живых клеток. Однако, как видно из этих обзоров, ряд важных моментов, связанных с ограничениями (например, существование жестких связей между переменными как следствие взаимодействия двух или более ограничений, существование излишних ограничений), не получили должного освещения. Как показала практика работы с биохимической стехиометрией, неучет вышеупомянутых свойств данной задачи приводит к потере решений.

Другой интересный момент состоит в том, что решения уравнений (11) с ограничениями (9) могут содержать слагаемые, которые по своему физическому смыслу являются излишними и отсутствуют в реальном объекте, так как последний имеет механизмы, препятствующие процессам, описываемым такими слагаемыми. Например, совокупность биохимических реакций, реализующая превращение заданных субстратов в заданные продукты и полученная как решение описываемой здесь алгебраической задачи, может содержать циклические цепочки реакций, которые в реальных клетках не функционируют благодаря действию регуляторных кинетических взаимодействий (последние не описываются системой стехиометрических уравнений). Пример — цепочка реакций, представляющая собой так называемый футильный цикл, баланс которой имеет вид $АТФ + Н_2О \rightarrow АДФ + Н_3РО_4$. Вариантов футильных циклов может быть сколь угодно много. Возможны также слагаемые, которые описывают не потоки вещества, а химические равновесия между метаболитами, несущественные для метаболического пути.

Выявить излишние слагаемые \vec{z} можно, решая задачу тем же способом, но при нулевых компонентах вектора \vec{b} для основной части задачи. Вычитание этого решения из полученного при $\vec{b} \neq 0$ позволяет убрать из картины метаболических потоков «паразитные» процессы. Для системы биохимических реакций это будет минимизация числа реакций, участвующих в искомом биохимическом пути. Такой подход в вычислительном отношении намного экономнее минимизации путем перебора реакций, исключаемых из первоначально полученного множества (см. цит. выше обзоры).

Выражаю благодарность к. б. н. М. А. Бабошину за помощь при поиске литературы, относящейся к приложениям методов линейной алгебры к стехиометрии метаболических путей.

Приложение 1. Перпендикуляр к гиперплоскости

Приведенные ниже рассуждения справедливы также для перпендикуляра к прямой или обычной плоскости. Для рассматриваемых в данной работе задач представляют интерес как уравнение прямой, перпендикулярной к вышеназванным геометрическим объектам, так и нахождение с ее помощью расстояния между ними и заданной точкой.

Пусть имеется гиперплоскость H , произвольно ориентированная в гиперпространстве. Ее уравнение:

$$\vec{h}\vec{y} = \sum_{i=1}^L h_i y_i = f, \quad (\text{П.1})$$

где \vec{h} — вектор нормали. Пусть, далее, имеется точка \vec{y}^{sp} , не лежащая на этой гиперплоскости. Расстоянием от \vec{y}^{sp} до H является длина перпендикуляра, опущенного из точки на гиперплоскость. Уравнение прямой, проходящей через \vec{y}^{sp} , запишем в виде

$$\vec{y} = \vec{y}^{sp} + \alpha \vec{k}, \quad (\text{П.2})$$

где α — параметр, \vec{k} — вектор, задающий направление прямой в гиперпространстве. Для перпендикуляра по отношению к гиперплоскости H в качестве \vec{k} можно взять вектор \vec{h} . Точ-

ка \vec{y}^{pr} пересечения прямой с гиперплоскостью должна удовлетворять и (П.2), и (П.1). Подставляя

$$\vec{y}^{pr} = \vec{y}^{sp} + \alpha_{\perp} \vec{h} \quad (\text{П.3})$$

в (П.1) в качестве \vec{y} , получаем значение параметра $\alpha = \alpha_{\perp}$, при котором имеет место пересечение:

$$\alpha_{\perp} = (f - \vec{h}\vec{y}^{sp}) / |\vec{h}|^2. \quad (\text{П.4})$$

Соответственно, найдено и значение \vec{y}^{pr} . Сопоставляя (П.4) с (17), видно, что если плоскость (П.1) является границей между положительными и отрицательными значениями x_n , то расположение \vec{y}^{sp} в области, соответствующей $x_n > 0$, соответствует $\alpha_{\perp} > 0$; если же точка \vec{y}_1 находится на самой этой плоскости, то $\alpha_{\perp} = 0$.

Вектор отрезка от \vec{y}^{pr} до \vec{y}^{sp} равен $\vec{P} = \vec{y}^{pr} - \vec{y}^{sp} = \alpha_{\perp} \vec{h}$. Квадрат расстояния от заданной точки до гиперплоскости (или плоскости, или прямой) равен

$$P^2 = (f - \vec{h}\vec{y}^{sp})^2 / |\vec{h}|^2 = \left(f - \sum_{i=1}^L h_i y_i^{sp} \right)^2 / \sum_{i=1}^L (h_i)^2. \quad (\text{П.5})$$

Приложение 2. Проекция вектора на пересечение гиперплоскостей

Рассмотрим группу пересекающихся гиперплоскостей:

$$\vec{h}_n \vec{y} = f_n, \quad (\text{П.6})$$

где n — номер гиперплоскости, число которых здесь обозначим как N^{hp} ; \vec{h}_n — нормаль к ней. Рассматривается случай, когда пересечение является связной областью, то есть не распадается на две и более отдельных области в пространстве $\{y\}$.

Спроецировать любой вектор, расположенный в пространстве $\{y\}$, на гиперплоскость можно, найдя проекцию на нее двух точек — начала и конца этого вектора. Обозначим координаты проецируемой точки как \vec{y}^{sp} , ее проекции — как \vec{y}^{pr} . Проецирование точки \vec{y}^{sp} на одну гиперплоскость H_n производится вдоль \vec{h}_n — вектора нормали к H_n , а координаты \vec{y}^{pr} находятся согласно (П.3) и (П.4).

Проецирование точки \vec{y}^{sp} на пересечение гиперплоскостей должно происходить вдоль вектора, являющегося линейной комбинацией нормалей \vec{h}_n . Наглядно это можно себе представить, рассматривая обычные плоскости в трехмерном пространстве, пересечением которых является прямая. Перпендикуляры к этим плоскостям и к этой прямой лежат в одной плоскости (перпендикулярной к первым двум). Поэтому перпендикуляр к прямой направлен вдоль линейной комбинации нормалей упомянутых двух плоскостей. Тогда в формуле (П.3) вместо \vec{h} должна стоять линейная комбинация $\sum_n \lambda_n \vec{h}_n$:

$$\vec{y}^{pr} = \vec{y}^{sp} + \alpha_{\perp} \sum_n \lambda_n \vec{h}_n,$$

а вместо (П.4) получаем систему уравнений

$$\alpha_{\perp} \sum_n \lambda_n \vec{h}_m \vec{h}_n = f_m - \vec{h}_m \vec{y}^{sp}, \quad m = 1, 2, \dots, N^{hp},$$

где $\vec{h}_m \vec{h}_n = \sum_l h_{ml} h_{nl}$ — квадратная матрица размера $N^{hp} \times N^{hp}$. Неизвестными здесь являются λ_m (их число равно N^{hp}) и α_{\perp} . Поскольку, однако, линейная комбинация векторов задает одно и то же направление в пространстве, если все ее коэффициенты умножить на одно и то же число, то все произведения $\alpha_{\perp} \lambda_n$ можно заменить на λ_n , и тогда имеем N^{hp} уравнений только для λ_n :

$$\sum_n \lambda_n \vec{h}_m \vec{h}_n = f_m - \vec{h}_m \vec{y}^{sp}, \quad m=1,2,\dots,N^{hp}. \quad (\text{П.7})$$

Соответственно,

$$\vec{y}^{pr} = \vec{y}^{sp} + \sum_n \lambda_n \vec{h}_n. \quad (\text{П.8})$$

Если $\text{rank } \vec{h}_m \vec{h}_n = N^{hp}$ (как и ранг расширенной матрицы системы (П.7)), то для заданной точки \vec{y}^{sp} существует единственный набор значений λ_n (зависящий от \vec{y}^{sp} !). Если $\text{rank } \vec{h}_m \vec{h}_n < N^{hp}$, то это означает, что одна или несколько гиперплоскостей не влияют на размерность области пересечения. (В вышеупомянутом трехмерном примере это случай, когда на одной и той же прямой пересекаются не две, а три или больше плоскостей.) Тогда одна или несколько нормалей должны быть удалены из системы (П.7) как лишние.

Если находятся проекции двух точек пространства $\{y\}$ на пересечение одной и той же совокупности гиперплоскостей, то система (П.7) будет записана для каждой из этих двух точек \vec{y}^{sp1} и \vec{y}^{sp2} . Пусть \vec{y}^{sp1} — координаты начала некоторого вектора \vec{g} , а \vec{y}^{sp2} — координаты его конца. Тогда $\vec{g} = \vec{y}^{sp2} - \vec{y}^{sp1}$. Коэффициенты λ_n будут иметь разные значения для \vec{y}^{sp1} и \vec{y}^{sp2} — соответственно $\lambda_n^{(1)}$ и $\lambda_n^{(2)}$. Взяв разность выражений (П.7) для \vec{y}^{sp2} и \vec{y}^{sp1} и обозначив $\lambda_n^{(2)} - \lambda_n^{(1)}$ как $\lambda_n^{(g)}$, получаем

$$\sum_n \lambda_n^{(g)} \vec{h}_m \vec{h}_n = -\vec{h}_m \vec{g}, \quad m=1,2,\dots,N^{hp}. \quad (\text{П.9})$$

Искомая проекция \vec{g}^{pr} вектора \vec{g} на пересечение гиперплоскостей, согласно (П.8), равна

$$\vec{g}^{pr} = \vec{g} + \sum_n \lambda_n^{(g)} \vec{h}_n. \quad (\text{П.10})$$

В качестве примера приведем трехмерный случай, когда \vec{g} проецируется на некоторую плоскость с нормалью \vec{h} . Тогда $N^{hp} = 1$, единственный коэффициент $\lambda^{(g)}$, согласно (П.9), равен $\lambda^{(g)} = -\frac{\vec{h}\vec{g}}{|\vec{h}|^2}$ и, согласно (П.10), $\vec{g}^{pr} = \vec{g} - \frac{(\vec{h}\vec{g})}{|\vec{h}|^2} \vec{h} = \vec{g} - \left(\frac{\vec{h}}{|\vec{h}|} \vec{g} \right) \frac{\vec{h}}{|\vec{h}|}$. Рассмотрение методом обычного векторного анализа дает следующее: проекция вектора \vec{g} на плоскость равна разности $\vec{g}^{pr} = \vec{g} - \vec{g}^{norm}$, где \vec{g}^{norm} — проекция \vec{g} на нормаль к плоскости. Единичный вектор, задающий направление нормали \vec{h} , равен $\frac{\vec{h}}{|\vec{h}|}$, модуль проекции \vec{g} на \vec{h} равен $\frac{\vec{h}}{|\vec{h}|} \vec{g}$, откуда следует приведенное здесь выражение для \vec{g}^{pr} , полученное как частный случай (П.10).

Список литературы

- Брандт Э.* Анализ данных. — М.: Мир, АСТ, 2003. — 686 с.
- Васильев Ф. П.* Методы оптимизации. Т. 1. — М.: Издательство МЦНМО, 2011. — 619 с.
- Рокафеллар Р.* Выпуклый анализ. — М.: Мир, 1973. — 468 с.
- Степанов Н. Ф., Ерлыкина М. Е., Филиппов Г. Г.* Методы линейной алгебры в физической химии. — М.: Издательство МГУ, 1976. — 360 с.
- Циглер Г. М.* Теория многогранников. — М.: Издательство МЦНМО, 2014. — 565 с.
- Buchanan B. B., Gruissem W., Jones R. L.* Biochemistry and Molecular Biology of Plants. — Rockville, American Society of Plant Physiologists, 2000. — 1367 p.
- Papin J. A., Stelling J., Price N. D., Klamt S., Schuster S., Palsson B. O.* Comparison of Network-Based Pathway Analysis Methods. Trends in Biotechnology. — 2004. — Vol. 22. — No. 8. — P. 400–405.
- Schilling C. H., Schuster S., Palsson B. O., Heinrich R.* Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-genomic Era // Biotechnology Progress. — 1999. — Vol. 15. — No. 3. — P. 296–303.