УДК 519.816
ББК 22.18

ЭКСПЕРИМЕНТАЛЬНОЕ СОПОСТАВЛЕНИЕ МЕТОДОВ ВЗВЕШЕННОЙ СУММЫ, ТЕОРИИ ПОЛЕЗНОСТИ И ТЕОРИИ ВАЖНОСТИ КРИТЕРИЕВ ДЛЯ РЕШЕНИЯ МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧ С БАЛЛЬНЫМИ КРИТЕРИЯМИ [1]

Салтыков С. А.[2]
(*Учреждение Российской академии наук Институт проблем управления РАН, Москва*)

Приводится экспериментальное сопоставление методов взвешенной суммы, теории полезности и теории важности критериев для решения многокритериальных задач с балльными критериями. Показано, что при весьма нестрогих допущениях метод взвешенной суммы приводит к существенному проценту ошибок, а методы теории полезности довольно часто «заставляют» «терять время» на построение функции полезности тогда, когда ее можно не строить. Теория важности критериев свободна от этих недостатков.

Ключевые слова: теория важности критериев, процент ошибок метода взвешенной суммы.

1. Введение

В целом ряде практических задач принятия решений варианты оцениваются по нескольким критериям, имеющим общую

---

[1] *Автор выражает благодарность проф. Подиновскому В. В. и д.ф.-м.н. Чеботареву П. Ю. за помощь в работе над статьей, ценные предложения, идеи и замечания.*

[2] *Салтыков Сергей Анатольевич, младший научный сотрудник* (*ssaltykov@mail.ru*)

балльную шкалу. Например, при отборе наиболее эффективного экспертного метода для решения данной прогнозной задачи [14] часто складывается такая ситуация, что один метод больше подходит по одним критериям (например, решение им задачи занимает меньше времени) и меньше подходит по другим (например, стоимость решения им задачи выше). Быстроту и стоимость решения можно оценивать по балльной шкале; также можно определить, как, с точки зрения ЛПР, важность быстроты решения соотносится с важностью стоимости решения. С другой стороны, если в исходной задаче критерии имеют разные шкалы, то их можно привести к единой балльной шкале [8].

Решать многокритериальные задачи можно разными методами: методом теории полезности (с использованием функции полезности, или ценности) [3], теории важности критериев [8], методом взвешенной суммы, взвешенного произведения (при общем подходе – взвешенного степенного среднего) и взвешенной медианы; также существует метод согласования кластеризованных ранжировок [2]. В данной работе мы проанализируем первые три метода. Итак, какой же из этих методов выбрать для решения многокритериальной задачи?

Считается, что необоснованное использование метода взвешенной суммы некорректно по многим причинам, в частности, из-за того, что внутрь этого метода «зашито» представление о том, что предпочтения ЛПР при переходе от одной шкальной градации к другой растут равномерно, что на практике далеко не всегда верно. Но это лишь теоретическое соображение: фактически это утверждение, что надо корректно использовать математическую теорию измерений и, в частности, методы работы с порядковыми шкалами. А насколько сильно некорректность использования операции взятия средневзвешенного – основы метода взвешенной суммы – для балльных оценок может отразиться на результатах анализа практических задач? Стоит ли её применять в условиях ограниченного времени (даже сознавая её некорректность)? И, соответственно, есть ли шанс, что данная стоящая перед нами многокритериальная задача будет решена верно? И если он есть, то насколько он велик? Достаточен ли он для практической значимости?

Аналогичный вопрос выбора возникает и при сопоставлении теории полезности и теории важности критериев. Можно ли получить какие-либо существенные выводы относительно предпочтительности выбора того или иного подхода? То есть не только лишь теоретические выводы, а количественные соображения, позволяющие создать основу для технико-экономического обоснования выбора наиболее приемлемого многокритериального подхода.

Для ответа на эти вопросы был проведен вычислительный эксперимент в рамках вероятностного подхода к анализу эффективности многокритериальных решающих правил [11]. В отличие от проводившихся ранее исследований [1, 12], в которых предполагалось, что информация о важности критериев фиксирована, мы приняли, что и оценки важности критериев являются случайными величинами. Это позволило получить оценки эффективности методов для всего массива возможных многокритериальных задач.

Кроме того, отметим интересные работы [4, 5], где также ставился вопрос о том, всегда ли ошибаются те, кто используют среднее арифметическое для порядковых величин, можно ли в какой-то мере «реабилитировать» для каких-то случаев такое его использование? В указанных работах был получен вывод о том, что можно при больших объемах выборок использовать среднее арифметическое различных значений двух каких-либо величин, измеренных в порядковой шкале, для определения, какая величина больше, если функции распределения этих величин не пересекаются, т. е. одна лежит над другой (но они могут касаться). «Последнее условие выполняется, если два значения некоторой величины измеряются с помощью одного и того же средства измерения, у которого распределение погрешностей не меняется при переходе от измерения одного значения рассматриваемой величины к измерению другого» [5].

Данная работа является продолжением и развитием исследования [13].

2. Описание используемой многокритериальной модели

Многокритериальная модель, рассматриваемая в статье, может быть представлена следующим образом:

(1) $<S, K_1, \ldots, K_m, R>$,

где S – это множество вариантов решений (стратегий, планов, альтернатив и т.д.), далее называемое множеством вариантов; $K_1, \ldots, K_m$ – критерии (целевые функции и т.д.); R – отношение нестрогого предпочтения.

Поясним представленную модель (1). Каждый вариант s из множества S всех (данных) вариантов характеризуется значениями $m \geq 2$ критериев K_i. Под критерием K_i мы понимаем функцию, определенную на множестве S и принимающую значения из множества X_i, называемого шкалой (а также множеством оценок, шкальных градаций, значений критериев). Без ограничения общности будем считать, что все оценки выражены в численном виде, и большие значения предпочтительней меньших. Таким образом, каждый вариант s характеризуется значениями $K_i(s)$ всех критериев, формирующих вектор оценок этого варианта, или его векторную оценку $x(s) = (K_1(s), \ldots, K_m(s))$. Следовательно, варианты сравниваются по предпочтительности посредством сопоставления их векторных оценок. Множество всех векторов оценок $X = X_1 \times \ldots \times X_m$. Мы предполагаем далее, что критерии являются однородными, т. е. имеют одинаковую (общую) шкалу $X_0 = X_1 = \ldots = X_m$ (так что $X = X_0^m$); более того, если критерий K_j заменить на $\xi(K_j)$, где ξ – некоторое допустимое преобразование, определяемое типом шкалы, то и все остальные критерии K_i следует заменить на $\xi(K_i)$. Примем также, что множество X_0 конечно: $X_0 = \{1, \ldots, q\}$. Элементы этого множества будем называть шкальными градациями.

Предпочтения ЛПР моделируются отношением предпочтения R на X: xRy означает, что вектор оценок x не менее предпочтителен, чем y. Отношение R порождает отношения безразличия I и (строгого) предпочтения P: xIy имеет место, когда справедливо xRy и yRx; xPy выполнено, когда верно xRy и неверно yRx.

В данной работе рассматриваются отношения нестрогого предпочтения Парето P^0; количественной важности критериев с порядковой шкалой R^{Θ} или шкалой первой порядковой метрики при замедлении роста предпочтений вдоль шкалы $R^{\Theta\&D}$ [9, 10], а также отношение предпочтения $R^{\Theta\&D^{||}}$, использующее интервальную информацию о скорости роста предпочтений вдоль шкалы критериев.

Информацию о количественной важности критериев будем использовать в форме значений важности критериев – чисел β_i, $1 \le i \le m$, так как это более удобно для целей проводимого вычислительного эксперимента. Функцию ценности будем задавать в аддитивном виде, сопоставляя каждой шкальной градации k её ценность $v(k)$.

3. Предварительные замечания

Метод взвешенной суммы, упомянутый во введении, является частным видом метода обобщенного критерия. Поэтому при характеристике метода взвешенной суммы, который на практике используется значительно чаще, чем методы теории полезности и теории важности критериев, надо обязательно отметить следующее относительно метода обобщенного критерия [6].

Согласно этому методу, все критерии вначале приводятся к сопоставимому виду («нормализуются»), т. е. заменяются на $K_i^o = \eta_i(K_i)$. И лишь затем строится обобщенный критерий $F(\alpha_1, K_1^0, ..., \alpha_m, K_m^0)$. Обоснованный выбор видов функций η_i и вида функции свертки F – очень сложная проблема, которой обычно уделяется мало внимания. Чаще всего попросту берут

$$\eta_i(K_i) = \frac{K_i - K_{i*}}{K_i^* - K_{i*}}$$

или

$$\eta_i(K_i) = \frac{K_i^* - K_i}{K_i^* - K_{i*}}, \qquad F = \sum_{i=1}^{m} \alpha_i K_i^0,$$

где K_i^* и K_{i*} – соответственно наибольшие и наименьшие значения критериев на множестве вариантов (первая формула для η_i применяется, если критерий K_i желательно максимизировать, вторая – минимизировать). С другой стороны, открытой остается проблема обоснованного назначения величин коэффициентов важности (относительных весов[1]) критериев α_i, так как строгого определения самого понятия важности критериев в методе обобщенного критерия нет.

В том случае, когда все критерии K_i изначально имеют общую шкалу и в роли обобщенного критерия используется взвешенная сумма $F=\sum_{i=1}^{m}\alpha_i K_i$, фактически принимается ничем не подкрепленное допущение о том, что рост предпочтений вдоль шкалы равномерен, а проблема обоснованного оценивания коэффициентов важности остается открытой.

Итак, проведем численные эксперименты для получения ответов на вопросы, поставленные во введении, последовательно привлекая все более «сильную» информацию об изменении предпочтений вдоль шкалы критериев. При этом везде будем полагать, что значения важности критериев известны.

Сначала исследуем случай, когда информации о скорости роста предпочтений вдоль шкалы критериев нет и, следовательно, можно применить методы теории количественной важности критериев с порядковой шкалой. То есть варианты в эксперименте будут сравниваться по отношению нестрогого предпочтения R^{Θ}.

Затем будем предполагать наличие информации о том, что предпочтения ЛПР растут вдоль шкалы (при движении от

---

[1] *Понятия относительного веса и коэффициента важности следует различать, но мы не будем на этом здесь останавливаться.*

меньших градаций к большим) с затуханием (так называемая информация D) и, следовательно, можно применить методы теории количественной важности критериев со шкалой первой порядковой метрики. То есть варианты в этом численном эксперименте будут сравниваться по отношению нестрогого предпочтения $R^{\Theta\&D}$.

Наконец, положим, что имеется информация о том, что рост предпочтений ЛПР не просто замедляется при движении от меньших шкальных градаций к большим, а «существенно» замедляется, т. е. зададимся определенными пределами для величины этого затухания. В данном случае варианты в эксперименте будут сравниваться по отношению предпочтения $P^{\Theta\&D^{||}}$.

Расчеты будут проводиться при помощи специально разработанной компьютерной программы. Для контроля полученных результатов будет применяться известная система DASS [7].

4. Описание вычислительного эксперимента с использованием отношения предпочтения R^{Θ} и его результатов

Вычислительный эксперимент строился следующим образом. Фиксировалась размерность многокритериальной задачи некоторой размерности (например, 7 вариантов и 5 критериев). Генератор случайных чисел выдавал балльные векторные оценки для каждого из вариантов. Баллы использовались от 2 до 5 по аналогии со школьными оценками, понятными и знакомыми большинству ЛПР.

Поясним, почему мы считаем достаточно репрезентативными выводы, полученные с использованием «урезанной» пятибалльной «школьной» шкалы. С математической точки зрения она не имеет никаких преимуществ перед другими шкалами, а вот с когнитивной – имеет. Можно потребовать эксперта/ЛПР интерпретировать полезность/ценность скорости, эффективности, массы или чего-то еще в общей шкале с любым числом градаций, но будет ли процесс этой интерпретации когнитивно корректным? Сможет ли человек протянуть ассоциативную цепочку в своих представлениях от весьма специализированных характеристик (скорости, красоты дизайна) к довольно общим – «удовлетворительно для меня», «хорошо», «очень хорошо»? Существует ли в его представлениях об исследуемых объектах и процессах такая ассоциативная цепочка? Кажется разумным, что необходимым условием существования этой цепочки является очень большая частота употребления общей шкалы к весьма разнородным феноменам. Так ли много таких общих шкал существует для нашей культурной подсистемы? По-видимому, совсем не много. Это школьная пятибалльная шкала и различные ее вариации («5–», «4+»), из которых можно сделать, скажем, весьма дифференцированную 12–13-балльную общую шкалу. Эти шкалы и являются самыми распространенными. В этом свете использование «школьной» пятибалльной шкалы представляется достаточно репрезентативным. Естественно, для других систем школьного образования «привычная» (а следовательно, так скажем, «когнитивно корректная») общая шкала может быть другой. Здесь видна культурная обусловленность количественных выводов вычислительного эксперимента и аналитических выкладок.

Появление любой из балльных оценок по каждому критерию было равновероятным. После этого генератором случайных чисел определялась важность критериев. Она определялась не через напрямую коэффициенты важности, а через N-модель [8, 9]. Если (для простоты интерпретации) представлять одну из многокритериальных задач, решаемых в эксперименте, задачей выбора наилучшего студента по его оценкам по разным предметам, тогда элемент N-модели может иметь следующий смысл: это число равноважных разделов в данном предмете. Предполагалось, что в одном предмете максимально содержится от 5 до 9 таких «разделов». Также предполагалось, что предмет может иметь любое число «разделов» от 1 до максимального с равной вероятностью. При решении многокритериальных задач одной размерности максимальное число разделов в предмете считалось постоянным. Таким образом, генератором случайных чисел определялись значения важности критериев, а не коэффициенты их важности.

Если удается выделить единственный недоминируемый вариант, т. е. лучший, (точнее, один недоминируемый класс эквивалентности вариантов), это значит, что информации только о балльных оценках и важности критериев для данной задачи достаточно. А следовательно, использование для этой задачи теории полезности (с построением функции полезности) было бы пустой тратой времени, так как для данной задачи выяснять, как изменяются предпочтения ЛПР вдоль шкалы, нет необходимости: как бы они ни изменялись (при любых их изменениях) с такими балльными оценками и значениями важности критериев у данных вариантов всегда будет лучшим тот вариант, который является недоминируемым по методу количественной важности критериев.

Если в рассматриваемой гипотетической задаче остался лишь один недоминируемый вариант (с точностью до эквивалентности), то это говорит и о том, что в данном случае метод взвешенной суммы дал бы верный ответ. И мы это можем утверждать, даже не зная, как изменяются предпочтения ЛПР вдоль шкалы [9].

Решался миллион таких задач одной размерности (случайно сгенерированных). И определялась доля случаев, когда для задачи данной размерности использование информации о количественной важности критериев достаточно для выделения одного лучшего (с точностью до эквивалентности) варианта.

Рассматривались и решались задачи с числом вариантов от 2 до 9 и числом критериев от 2 до 9. Результаты расчетов занесены в таблицу 1. В этой таблице, а также в таблицах 2–5, для облегчения поиска цветом выделены числа, к которым есть отсылки в тексте статьи. Таким образом, в таблицу занесено 64 числа, что соответствует 64 миллионам решенных задач. Все эти 64 миллиона задач решались в предположении, что существует 5 различных значений важности критериев (говоря языком нашего иллюстративного примера, в предмете может быть от 1 до 5 равноважных разделов).

Таблица 1. Доля случаев, когда количественной важности критериев достаточно для определения одного (с точностью до эквивалентности) недоминируемого варианта; 5 различных значений важности критериев

Число вариантов	Число критериев							
	2	3	4	5	6	7	8	9
2	0,850	0,738	0,669	0,627	0,602	0,585	0,574	0,566
3	0,778	0,624	0,536	0,484	0,454	0,435	0,422	0,415
4	0,754	0,574	0,472	0,413	0,380	0,359	0,345	0,336
5	0,753	0,554	0,439	0,374	0,336	0,313	0,298	0,287
6	0,763	0,549	0,422	0,350	0,307	0,281	0,265	0,255
7	0,778	0,550	0,412	0,333	0,288	0,259	0,243	0,231
8	0,796	0,554	0,406	0,323	0,273	0,243	0,224	0,212
9	0,813	0,560	0,404	0,315	0,263	0,230	0,210	0,197

Далее были проведены аналогичные эксперименты для 6, 7, 8, 9 различных значений важности критериев. Полученные в этих исследований результаты не сильно отличаются от приведенных в таблице 1, и эти различия не влияют на выводы, получаемые в статье. Поэтому далее использовались 5 различных значений важности критериев.

Для повышения надежности результаты, представленные в таблице 1, были получены двумя разными путями, т. е. одна и та же многокритериальная задача решалась двумя способами – алгоритмом сравнения по отношению предпочтения R^{Θ} с построением N-модели и без построения оной [9].

Если во множестве вариантов присутствует, говоря образно, слишком много «хороших» вариантов, «сильно» превосходящих остальные варианты, то это выглядит несколько неправдоподобно. В действительности такое бывает не часто. Обычно одни варианты, среди которых производится отбор, лучше прочих по одним параметрам и хуже по другим.

Для того чтобы отразить этот факт в вычислительном эксперименте, на варианты, создаваемые генератором случайных чисел, накладываем следующее ограничение: сумма балльных оценок в векторе, соответствующем варианту, не должна превышать 4, умноженное на число оценок в варианте (что равно числу критериев). Это следует из такого соображения: никакая оценка, доминирующая по Парето, над оценкой из всех «четверок», не должна удовлетворять этому ограничению и, следовательно, не должна попасть в экспериментальное множество вариантов:

$$(2)\quad \sum_{i=1}^{m} K_i(s) \le 4m.$$

Кроме ограничения (2) на множестве вариантов S, рассмотрим, исходя из тех же соображений, также и следующее ограничение: сумма квадратов балльных оценок не должна превышать число 16, умноженное на число оценок в варианте:

$$(3)\quad \sum_{i=1}^{m} K_i^2(s) \le 16m.$$

На рис. 1 изображено множество векторов оценок для случая двух критериев ($m = 2$), а также наложенные ограничения (2) и (3). Варианты, изображенные на рисунке черным, не удовлетворяют ни ограничению (2), ни (3). Варианты, изображенные серым, удовлетворяют (2), но не удовлетворяют (3).

При наложенном ограничении (2), говоря неформально, «отрезаются» только «очень хорошие» варианты, а при ограничении (3) – «очень хорошие» и «хорошие». Поэтому понятно, почему результаты эксперимента с ограничением (2) находятся между результатами эксперимента без наложенного ограничения и с ограничением (3). Поэтому мы предположили, что наиболее существенные выводы эксперимента лежат между результатами эксперимента без наложенного ограничения и с ограничением (3).

Отметим еще, что во всех случаях вероятность того, что метод взвешенной суммы даст правильный ответ, выше, чем равномерный случайный выбор из множества вариантов. Из таблицы 1 видно, что любое число в этой таблице больше, чем единица, деленная на соответствующее число вариантов, соответствующих этому числу в ячейке. Например, для 7 вариантов и 9 критериев: 0,231 > 1/9.

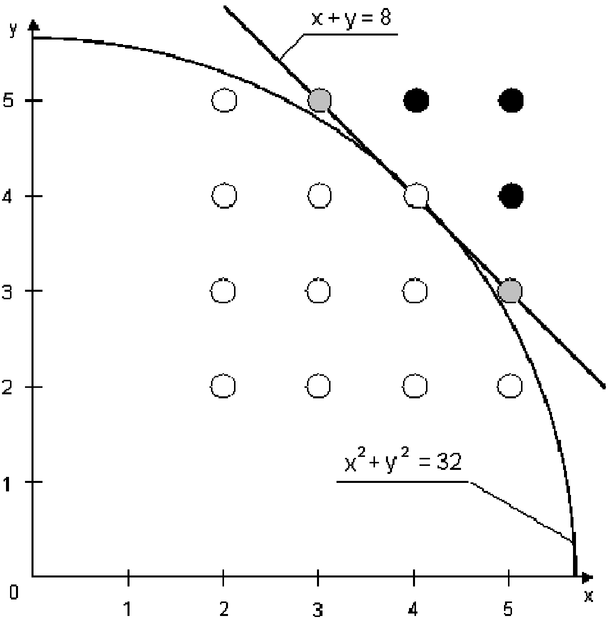

Рис. 1. Множество векторов оценок для случая двух критериев

5. Описание вычислительного эксперимента с использованием отношения предпочтения $P^{\Theta\&D}$ и анализ результатов

По аналогичной схеме проводились исследования для случая, когда имеется еще и информация *D* о том, что рост предпочтений ЛПР замедляется, «затухает» вдоль шкалы; результаты эксперимента занесены в таблицу 2. Также проведен эксперимент при условии, что на множество вариантов наложено ограничение (3), результаты занесены в таблицу 3.

По полученным результатам видно, что в большинстве случаев удается определить один недоминируемый вариант (с точностью до эквивалентности) при использовании отношения предпочтения $P^{\Theta\&D}$. Также видим, что эта доля случаев существенно выше, чем при использовании отношения P^{Θ}. Таким образом, получается, что добавление информации D о том, что предпочтения вдоль шкалы растут замедленно, является существенным, ее влияние оказывается очень сильным, а использование – оправданным.

С точки зрения главных целей проводимого исследования результаты получаются еще более интересными: если предположить, что ситуация, описываемая информацией D, достаточно типична, то получается, что в построении функции полезности нет необходимости в очень существенной доле случаев.

Также отметим, что и здесь каждая экспериментальная многокритериальная задача решалась двумя способами, приведенными в [10]. Сделано это для повышения достоверности существенных выводов статьи.

Таблица 2. Доля случаев, когда количественной важности критериев и информации D достаточно для определения одного (с точностью до эквивалентности) недоминируемого варианта; 5 различных значений важности критериев

Число вариантов	Число критериев							
	2	3	4	5	6	7	8	9
2	0,938	0,888	0,863	0,832	0,813	0,816	0,806	0,807
3	0,903	0,831	0,783	0,743	0,731	0,720	0,710	0,713
4	0,898	0,801	0,739	0,698	0,660	0,649	0,637	0,633
5	0,890	0,788	0,705	0,659	0,635	0,609	0,594	0,585
6	0,897	0,773	0,696	0,637	0,603	0,581	0,551	0,558
7	0,907	0,778	0,691	0,628	0,589	0,564	0,535	0,530
8	0,910	0,787	0,694	0,618	0,575	0,539	0,517	0,508
9	0,920	0,781	0,683	0,605	0,550	0,533	0,501	0,481

Таблица 3. Доля случаев, когда количественной важности критериев и информации D достаточно для определения одного (с точностью до эквивалентности) недоминируемого варианта при ограничении (3) на S; 5 различных значений важности критериев

Число вариантов	Число критериев							
	2	3	4	5	6	7	8	9
2	0,908	0,862	0,823	0,810	0,803	0,792	0,788	0,779
3	0,833	0,749	0,705	0,690	0,692	0,680	0,671	0,662
4	0,778	0,676	0,623	0,606	0,610	0,603	0,592	0,586
5	0,732	0,616	0,552	0,542	0,548	0,552	0,529	0,538
6	0,719	0,571	0,510	0,511	0,498	0,499	0,494	0,490
7	0,704	0,527	0,471	0,455	0,460	0,464	0,449	0,458
8	0,680	0,500	0,442	0,420	0,433	0,437	0,430	0,427
9	0,672	0,477	0,413	0,393	0,398	0,410	0,407	0,400

6. Схема расчета вероятности ошибки при использовании метода взвешенной суммы и анализ полученных результатов

Теперь найдем нижнюю оценку (оценку снизу) вероятности того, что метод взвешенной суммы приведет к ошибке при определении наилучшего варианта (а точнее — класса эквивалентности таких вариантов). Ошибка заведомо может возникнуть тогда, когда

1) отношение $P^{\Theta\&D}$ выделяет один недоминируемый класс эквивалентности;

2) существуют как минимум два варианта, средневзвешенные которых равны между собой и больше средневзвешенных других вариантов;

3) среди вариантов, имеющих наибольшие средневзвешенные, есть те, которых нет в недоминируемом по $P^{\Theta\&D}$ классе эквивалентности.

Результаты расчетов для случаев, когда нет ограничения (3) и когда оно наложено на множество *S*, представлены в таблицах 4 и 5 соответственно.

Таблица 4. Проценты ошибок, которые делает метод взвешенной суммы, полученные при сравнении по отношению предпочтения $P^{\Theta\&D}$

Число вариантов	Число критериев							
	2	3	4	5	6	7	8	9
2	1,120	1,510	1,490	1,500	1,280	1,100	1,070	1,130
3	1,710	2,280	2,120	1,890	1,700	1,600	1,340	1,430
4	1,880	2,480	2,500	2,280	2,180	1,990	1,670	1,600
5	1,840	2,770	2,710	2,270	2,100	1,940	1,810	1,500
6	2,090	3,200	3,030	2,660	2,310	1,930	1,810	1,640
7	1,760	3,280	3,110	2,660	2,400	2,180	1,820	1,850
8	1,950	3,170	3,170	2,670	2,400	2,200	1,860	1,600
9	1,740	3,250	3,310	2,830	2,600	2,260	2,040	1,740

Таблица 5. Проценты ошибок, которые делает метод взвешенной суммы, полученные при сравнении по отношению предпочтения $P^{\Theta\&D}$ с ограничением (3) на S

Число вариантов	Число критериев							
	2	3	4	5	6	7	8	9
2	1,720	2,050	1,940	1,890	1,680	1,520	1,410	1,280
3	2,920	3,190	2,890	2,540	2,360	1,910	1,770	1,790
4	3,840	4,040	3,340	3,170	2,770	2,360	2,250	1,790
5	4,360	4,400	3,690	3,100	2,960	2,540	2,340	2,230
6	4,940	4,730	4,140	3,550	3,030	2,600	2,550	2,140
7	5,280	5,440	4,560	3,610	3,150	2,490	2,550	2,290
8	5,450	5,750	4,350	3,620	2,950	2,790	2,580	2,340
9	6,030	5,910	4,240	3,640	3,030	2,660	2,530	2,310

Данные в таблице 4 поясним на примере задачи размерности 7 вариантов на 7 критериев (ограничения на множество вариантов S не наложено). Мы знаем, что ошибка при использовании метода взвешенной суммы точно будет совершена в 2,18% случаев (см. таблицу 4). Мы знаем, что ошибка не будет совершена в $56{,}4 - 2{,}18 = 54{,}22$% случаев (56,4% – из таблицы 2). Подчеркнем, что числа в таблицах 4 и 5 – нижние границы для средних величин ошибок! Для установления более точных границ необходима более точная, в частности, интервальная информация относительно скорости роста предпочтений ЛПР вдоль шкалы.

7. Случай «существенного» убывания роста предпочтений вдоль шкалы, большого числа вариантов и малого числа критериев

Внимательное изучение предыдущих таблиц позволяет сделать вывод о том, что стоит отдельно рассмотреть случай решения многокритериальных задач, когда вариантов много, критериев мало и рост предпочтений ЛПР «существенно» убывает вдоль шкалы.

Оказывается, можно из простых соображений вывести формулу, позволяющую определить процент ошибок, получающихся при использовании метода взвешенной суммы для этого случая. Предположение, что вариантов много, позволяет сделать аналитический вывод этой формулы и не проводить для данного случая вычислительного эксперимента. Идея такова, что при малом числе градаций шкалы и числе критериев количество вариантов (разумеется, с разными векторными оценками) «технически» ограничено. Действительно, общее число всех векторных оценок есть q^m (в случае, если есть все градации от 1 до q). А раз так, значит, при достаточно большом числе вариантов во множество S попадут все возможные q^m вариантов с равной вероятностью. Предполагаем, что на множество S наложено ограничение (3). Тогда при четырех различных градациях и $m = 2$ различных возможных вариантов 16. С учетом наложенного на множество вариантов ограничения (3), видим, что при данных условиях существуют три варианта, недоминируемых по Парето (это, кстати, наглядно видно на рис. 1). Таким образом, решим следующую задачу отбора из этих трех недоминируемых по Парето вариантов одного недоминируемого (с точностью до эквивалентности) по отношению $R^{\Theta\&D^{||}}$ при различных значениях важности критериев β_1 и β_2.

$S_0 = \{s_1, s_2, s_3\}$,

(4) $s_1 = (5, 2)$, $s_2 = (4, 4)$, $s_3 = (2, 5)$.

При $\beta_1 = \beta_2$ можно показать, что единственным недоминируемым вариантов является s_2. Выведем формулу для определения величины ошибки при использовании метода взвешенной суммы, если значения важности критериев β_1 и β_2 не равны между собой. Без ограничения общности положим, что $\beta_1 > \beta_2$. Тогда

$\acute{\Omega} = \{1 \succ 2\}$,

$s_1 P^{\acute{\Omega}} s_3$,

$S_{\Omega} = \{s_1, s_2\}$.

Более того, можно показать, что добавление информации о количественной важности критериев, а также и информации D (о том, что предпочтения растут с «затуханием» вдоль шкалы, т. е. шкала является первой порядковой метрики) не уменьшают множество недоминируемых вариантов:

$S_{\Theta} = \{s_1, s_2\}$,

$S_{\Theta\&D} = \{s_1, s_2\}$.

Значит, придется совершенствовать шкалу, т. е. уточнять, как изменяются предпочтения ЛПР вдоль шкалы. Сделаем это по аналогии с [6]. Ограничения на β_1 и β_2 здесь не налагаются. Ценность (полезность) $v(j)$ градаций j шкалы неизвестна (т. е. функция ценности v не задана). Изначально лишь известно, что предпочтения возрастают вдоль шкалы, то есть $v(j) > v(k)$ при $j > k$ (иначе говоря, шкала критериев – порядковая). Далее предполагается также, что имеется дополнительная интервальная информация об убывающей скорости роста предпочтений вдоль шкалы критериев, заданная в виде неравенств:

$$d_k \le \frac{v(k+1)-v(k)}{v(k+2)-v(k+1)} \le u_k, \quad k = 1, ..., q-2.$$

Предположим для простоты расчета, что скорость роста предпочтений ЛПР убывает так, что $d_k = u_k$. Выведем формулу, показывающую величину отношений разности ценности градаций $d_k^0 = u_k^0$, при которой два оставшихся недоминируемых варианта s_1 и s_2 одинаковы по предпочтительности по отношению $P^{\Theta\&D^{||}}$ (третий, напомним, отбрасывается при сравнении по качественной важности). Соответственно, если относительно этого числа равномерный случай (предполагаемый методом взвешенной суммы, при котором $d_k = u_k = 1$) и неравномерный случай ($d_k = u_k > 1$) оказываются «по разные стороны», то решение методом взвешенной суммы приведет к ошибке. То есть, если это число $d_k^0 = u_k^0 = 1{,}3028$, а $d_k = u_k = 1{,}5$, то использование метода взвешенной суммы приведет к ошибке, так как $1 < 1{,}3028 < 1{,}5$.

Легко видеть, что s_1 и s_2 будут эквивалентны по важности, если $\beta_1 \times (v(5) - v(4)) = \beta_2 \times (v(4) - v(2))$.

Предположим, что отношения разности ценности градаций постоянны для всех градаций:

$$\frac{v(4)-v(3)}{v(3)-v(2)} = \frac{v(5)-v(4)}{v(4)-v(3)} = a.$$

Пусть

$v(3) - v(2) = 1,$

тогда

$v(4) - v(3) = a,$

$v(5) - v(4) = a^2,$

$$v(4) - v(2) = v(4) - v(3) + v(3) - v(2) =$$
$$= (v(4) - v(3)) + (v(3) - v(2)) = a + 1$$

отсюда:

$\beta_1 \times a^2 = \beta_2 \times (a + 1).$

У этого уравнения есть только один положительный корень:

$$a = \frac{\beta_2 + \sqrt{\beta_2^2 + 4\beta_1\beta_2}}{2\beta_1}.$$

Для него

$$d_k^0 = u_k^0 = (v(3) - v(2))/(v(4) - v(3)) = 1/a =$$

$$= \frac{2\beta_1}{\beta_2 + \sqrt{\beta_2^2 + 4\beta_1\beta_2}}.$$

Последняя формула для данных значений коэффициентов важности β_1 и β_2 (и данных вариантов s_1 и s_2, естественно) дает такое значение «затухания» роста предпочтений ЛПР, при котором варианты становятся эквивалентными. («Затухание» меньше единицы может быть истолковано как «возрастание скорости роста предпочтений».) Результаты для 5 различных значений важности критериев занесем в таблицу 6.

Таблица 6. Значения отношений разностей шкальных градаций, при которых варианты s_1 и s_2 в (4) эквивалентны

Значение важности второго критерия β_2 в (4)	Значение важности первого критерия β_1 в (4)				
	1	2	3	4	5
1	0,6180	1,0000	1,3028	1,5616	1,7913
2	1,0000	0,6180	0,8229	1,0000	1,1583
3	1,3028	0,8229	0,6180	0,7583	0,8844
4	1,5616	1,0000	0,7583	0,6180	0,7247
5	1,7913	1,1583	0,8844	0,7247	0,6180

Легко видеть, что если в таблице 6 число больше 1, то использование метода взвешенной суммы может привести к ошибке (если рост предпочтений затухает достаточно сильно). Если меньше единицы – не приведет. Если равно единице –

приведет к ошибке в половине случаев, условно предположим, что «неудачная» для метода взвешенной суммы – нижняя часть таблицы (под диагональю).

Из таблицы 6 видно, что в 2 случаях из 25 решение задачи методом взвешенной суммы приводит к ошибке в таком случае, что для того, чтобы это констатировать, достаточно информации D. Важно отметить, что эти 2/25 («две двадцать пятых»), то есть 0,08 (8%), полученные таким образом без вычислительного эксперимента, очень точно совпадают с теми же 8%, полученными в вычислительном эксперименте, когда число вариантов становится больше примерно 50.

Еще 8 случаев ошибок, совершаемых исследователем при использовании метода взвешенной суммы, добавляется, если известно, что рост предпочтений не просто «затухает», а «достаточно хорошо затухает». Так, если известно, что отношения разности ценностей градаций больше 1,8, то вероятность ошибки метода взвешенной суммы равна 10/25, то есть 40%. Эти десять «ячеек» таблицы, соответствующих ситуациям «с ошибками», в таблице выделены цветом.

Также из вышеприведенной таблицы видно следующее. Если отношения разностей ценностей градации

- от 1 до 1,1583, то вероятность ошибки 8%;
- от 1,1583 до 1,3027 – 16%;
- от 1,3027 до 1,5615 – 24%;
- от 1,5615 до 1,7913 – 32%;
- больше 1,7913 – 40%.

Для случая, когда выделяется 10 различных значений важности критериев, результаты занесены в таблицу 7.

Таблица 7. Значения отношений разностей шкальных градаций, при которых варианты s_1 и s_2 в задаче (4) эквивалентны

β_2	β_1									
	1	2	3	4	5	6	7	8	9	10
1	0,618	1,000	1,303	1,562	1,791	2,000	2,193	2,372	2,541	2,702
2	1,000	0,618	0,823	1,000	1,158	1,303	1,437	1,562	1,679	1,791
3	1,303	0,823	0,618	0,758	0,884	1,000	1,107	1,208	1,303	1,393
4	1,562	1,000	0,758	0,618	0,725	0,823	0,914	1,000	1,081	1,158
5	1,791	1,158	0,884	0,725	0,618	0,704	0,785	0,860	0,932	1,000
6	2,000	1,303	1,000	0,823	0,704	0,618	0,690	0,758	0,823	0,884
7	2,193	1,437	1,107	0,914	0,785	0,690	0,618	0,680	0,739	0,796
8	2,372	1,562	1,208	1,000	0,860	0,758	0,680	0,618	0,673	0,725
9	2,541	1,679	1,303	1,081	0,932	0,823	0,739	0,673	0,618	0,667
10	2,702	1,791	1,393	1,158	1,000	0,884	0,796	0,725	0,667	0,618

Здесь ошибка достигает 45/100 = 45%. Это видно наглядно из таблицы 7: 45 «ячеек» из 100, выделенные цветом, соответствуют ситуациям, когда использование метода взвешенной суммы приведет к ошибкам.

А если, скажем, предположить, что отношения разности ценности градаций лежит в довольно разумных пределах, скажем, [1,8; 2,5] для случая 10 различных значений важности критериев доля ошибок будет 35%, это наглядно (по клеточкам) можно посчитать по результатам, отображенным в таблице 8.

Таблица 8. Значения отношений разностей шкальных градаций, при которых варианты s_1 и s_2 в задаче (4) эквивалентны

β_2	β_1									
	1	2	3	4	5	6	7	8	9	10
1	0,618	1,000	1,303	1,562	1,791	2,000	2,193	2,372	2,541	2,702
2	1,000	0,618	0,823	1,000	1,158	1,303	1,437	1,562	1,679	1,791
3	1,303	0,823	0,618	0,758	0,884	1,000	1,107	1,208	1,303	1,393
4	1,562	1,000	0,758	0,618	0,725	0,823	0,914	1,000	1,081	1,158
5	1,791	1,158	0,884	0,725	0,618	0,704	0,785	0,860	0,932	1,000
6	2,000	1,303	1,000	0,823	0,704	0,618	0,690	0,758	0,823	0,884
7	2,193	1,437	1,107	0,914	0,785	0,690	0,618	0,680	0,739	0,796
8	2,372	1,562	1,208	1,000	0,860	0,758	0,680	0,618	0,673	0,725
9	2,541	1,679	1,303	1,081	0,932	0,823	0,739	0,673	0,618	0,667
10	2,702	1,791	1,393	1,158	1,000	0,884	0,796	0,725	0,667	0,618

Последний вывод, пожалуй, самый сильный: получается, что при разумном «затухании» роста предпочтений (существующем и заметном, но не чрезмерном) доля ошибок использования метода взвешенной суммы будет так велика – около 35%, а для 5 различных значений важности критериев – даже 40%. В заключение отметим следующее:

1. Численный эксперимент показывает, что при большом числе вариантов при трех критериях ошибка при использовании метода взвешенной суммы еще больше, чем при двух (9% против 8%) – наибольшая из всех случаев с числом критериев от 2 до 9, поэтому, по-видимому, имеет смысл то же сделать (т. е. аналитически вывести формулу для процента ошибок метода взвешенной суммы) и для случая 3 критериев, но из-за гораздо меньшей наглядности мы в этой работе не будем этого делать.

2. Информация *D* оказалась существенна для вывода об ошибках, а информация «о существенном затухании» – еще более существенна, она смогла увеличить долю ошибок в некотором классе ситуаций при разумных допущениях с 8–9% до 35–40%. По всей видимости, для случая большого числа критериев и/или малого числа вариантов информация о «существенном затухании» скорости роста предпочтений ЛПР тоже может оказаться очень полезной для более точного определения процента ошибок, совершаемых при применении метода взвешенной суммы по сравнению только лишь с «информацией *D*»; и там тоже эта доля может существенно увеличиться. Такое исследование можно провести позже. Возможно также проведение расчетов для случая наличия интервальных оценок скорости роста предпочтений при помощи известной системы DASS [7].

8. Выводы

Для многокритериальных задач с однородными критериями с числом вариантов от 5 до 9 и числом критериев от 5 до 9:

1. Теория полезности «заставляет терять время» в 39–66% случаев[1]: целесообразно вначале использовать теорию важности критериев, а не пытаться сразу же строить функцию полезности.

2. Метод взвешенной суммы делает ошибку не менее чем в 1,5–3,6 % случаев[2] при имеющейся информации о том, что рост предпочтений вдоль шкалы замедляется.

3. При наличии информации «о существенном затухании» роста предпочтений вдоль шкалы использование метода взвешенной суммы приведет к ошибке в некотором классе особо «спорных» ситуаций не менее чем в 35–40% случаев[3].

4. В экстренных случаях, когда по каким-либо причинам использовать ТВК невозможно, имеет смысл применить метод взвешенной суммы: вероятность того, что он даст правильный

---

[1] *39% - из таблицы 3; 66% – из таблицы 2.*
[2] *1,5% – из таблицы 4; 3,6% – из таблицы 5.*
[3] *35% – посчитано по таблице 8; 40% – по таблице 6.*

ответ заведомо выше, чем то, что правильный ответ будет получен, если один вариант из всего их множества будет выбран абсолютно случайно (т. е. «наобум»).

Вышеприведенные выводы приводят к самому, наверное, главному заключению: использование теории важности критериев целесообразно, она является практическим инструментом с оцениваемой эффективностью, а не просто неким интересным теоретическим построением.

Проведенное исследование позволяет проводить технико-экономическое обоснование применения теории важности критериев как эффективного инструмента анализа многокритериальных задач принятия решений и, таким образом, открывает для неё дверь в сферу инноваций.

Литература

1. БАРЫШНИКОВ Ю. М. *О среднем числе вариантов, недоминируемых по сравнению В. В. Подиновского* // Автоматика и телемеханика. – 1990. – №6. – С. 161–167.
2. ГОРСКИЙ В. Г., ГРИЦЕНКО А. А., ОРЛОВ А. И. *Метод согласования кластеризованных ранжировок* // Автоматика и телемеханика. – 2000. – №.3. – С.179–187.
3. КИНИ Р. Л., РАЙФА Х. *Принятие решений при многих критериях: предпочтения и замещения* / Пер. с англ. – М.: Радио и связь, 1981.
4. ОРЛОВ А. И. *Устойчивость в социально-экономических моделях*. – М.: Наука, 1979.
5. ОРЛОВ А. И. *Эконометрика*. – М.: Экзамен, 2002.
6. ПОДИНОВСКИЙ В. В. *Аксиоматическое решение проблемы оценки важности критериев в многокритериальных задачах принятия решений* // Современное состояние теории исследования операций. – М.: Наука, 1979. – С.117–145.
7. ПОДИНОВСКИЙ В. В. *Анализ задач многокритериального выбора методами теории важности критериев при помощи компьютерных систем поддержки принятия решений* //

Изв. АН. Теория и системы управления. – 2008. – № 2. – С. 64–68.
8. ПОДИНОВСКИЙ В. В. *Введение в теорию важности критериев в многокритериальных задачах принятия решений*. – М.: Физматлит, 2007.
9. ПОДИНОВСКИЙ В. В. *Количественная важность критериев* // Автоматика и телемеханика. – 2000. – № 5. – С. 110–123.
10. ПОДИНОВСКИЙ В. В. *Количественная важность критериев с дискретной шкалой первой порядковой метрики* // Автоматика и телемеханика. – 2004. – №8. – С. 196–203.
11. ПОДИНОВСКИЙ В. В. *Об оценке эффективности решающих правил в многокритериальных задачах* // Изв. АН СССР. Техн. кибернетика. – 1987. – № 1. – С. 3–9.
12. ПОДИНОВСКИЙ В. В. *Оценка эффективности решающих правил в дискретных многокритериальных задачах* // Методы оптимизации в экономико-математическом моделировании. – М.: Наука, 1991. – С. 308–324.
13. САЛТЫКОВ С. А. *Экспериментальное сопоставление различных многокритериальных подходов* // Материалы XXXVI Международной конференции «Информационные технологии в науке, социологии, экономике и бизнесе», Ялта, Гурзуф, май 2009 г. Приложение к журналу «Открытое образование». – С. 315–317.
14. СИДЕЛЬНИКОВ Ю.В., САЛТЫКОВ С.А. *Процедура установления соответствия между задачей и методом* // Экономические стратегии. – 2008. – №7. – С. 102–109.

EXPERIMENTAL COMPARISON OF METHODS OF WEIGHTED SUM, UTILITY THEORY AND CRITERIA IMPORTANCE THEORY FOR THE SOLUTION OF MULTICRITERIAL PROBLEMS WITH SCORE CRITERIA

Sergey Saltykov, Institute of Control Sciences of RAS, Moscow, junior researcher, (ssaltykov@mail.ru).

Abstract: An experimental comparison of the methods of weighted sum, the utility theory and the criteria importance theory is performed while solving the multicriteria problems with score criteria. It is shown that under very weak assumptions, the method of weighted sum leads to significant percentage of mistakes, while methods of the utility theory are often result in loss of time for unneeded construction of the utility function. The criteria importance theory is free of these shortcomings.

Keywords: criteria importance theory, percentage of errors in the method of weighted sum.

Статья представлена к публикации членом редакционной коллегии А. И. Орловым