ГЛОБАЛЬНЫЕ РЫНКИ
И ФИНАНСОВЫЙ ИНЖИНИРИНГ
Том 4 ● Номер 3 ● июль-сентябрь 2017
ISSN 2410-8618
Global Markets and Financial Engineering

издательство
Креативная экономика

Агрегация кредитных рейтингов как задача построения консенсуса в системе экспертных оценок

Буздалин А.В. [1], Заночкин А.Ю. [1], Курбангалеев М.З. [2], Смирнов С.Н. [2, 3]

[1] ЗАО «Интерфакс», Москва, Россия
[2] Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
[3] Московский государственный университет им. М.В. Ломоносова, Москва, Россия

АННОТАЦИЯ:

Информация о рейтингах компаний, присваиваемых кредитными рейтинговыми агентствами (КРА), рассматривается в статье как ранжирование компаний по относительному кредитному качеству. С учетом специфики задачи строится «справедливый» (в определенном смысле) способ агрегирования этой информации в «консенсусное» ранжирование всех компаний. Процедура применяется к реальным данным о рейтингах, присваиваемых КРА российским банкам в период 2010–2016 годов в национальных шкалах. Расчеты показывают, что полученное консенсусное ранжирование обладает высокой устойчивостью, а также может быть с успехом использовано в качестве объясняющей (скоринговой) переменной для оценки кредитного риска компаний. Предложенный способ является весьма общим и позволяет агрегировать ранжирования произвольной природы (рейтинги КРА, внутренние экспертные оценки, рыночные показатели), а также неполные ранжирования (например, когда компании рейтингуются только частью КРА или не имеют рыночных долговых инструментов).

КЛЮЧЕВЫЕ СЛОВА: кредитные рейтинги, экспертная информация, консенсусное ранжирование, медиана Кемени, лексикографическая оптимизация

Aggregation of credit ratings as a task of building consensus in the system of expert assessments

Buzdalin A.V. [1], Zanochkin A.Yu. [1], Kurbangaleev M.Z. [2], Smirnov S.N. [2, 3]

[1] Interfax JSC, Russia
[2] The National Research University Higher School of Economics (HSE), Russia
[3] Lomonosov Moscow State University, Russia

Введение

В кредитном анализе для оценки кредитного качества заемщиков используется множество показателей. Эти показатели обладают различной природой: показатели финансовой отчетности и рыночная информация (например, спрэды облигаций), профессиональные суждения, ранговые показатели (рейтинги) и количественные оценки (модельные вероятности дефолтов). Однако все эти показатели или их численные агрегаты независимо от природы можно рассматривать как объясняющие (скоринговые) переменные, которые порождают некоторые ранжирования[1] заемщиков по степени кредитного риска (относительному кредитному качеству).

Качество такого ранжирования определяется его способностью разделить «надежных» заемщиков и «потенциальных дефолтеров». Общепринятыми критериями разделяющей способности ранжирования являются CAP-(ROC-)кривые и показатели AR, которые[2] в том числе предусмотрены требованиями «Базеля II» (см. [2]). Чтобы повысить качество ранжирования, индивидуальные показатели принято агрегировать. На практике наиболее популярны[3] такие инструменты агрегирования, как логистическая регрессия (параметрический метод) и дерево решений (непараметрический метод).

[1] Ранжирование объектов здесь и далее будем понимать как нестрогое, т.е. возможно, что несколько объектов занимают в ранжировании одинаковое положение (в отличие от строгого ранжирования).

[2] CAP (Cumulative Accuracy Plot) и ROC (Receiver Operating Characteristic) представляют собой кривые Лоренца и являются графическими представлениями разделяющей способности дискриминационной переменной. AR (Accuracy Ratio) представляет собой коэффициент Джини, является численной характеристикой разделяющей способности и может быть определен как для CAP-, так и ROC-кривых. Подробнее см. [1] *(Tasche, 2010)*.

[3] Предостережем от некритического подхода при агрегации: зависимость кредитного качества конкретного показателя может быть нелинейной и немонотонной.

ABSTRACT:

Information on companies' ratings assigned by credit rating agencies (CRA) is considered in the article as a ranking of companies by relative credit quality. Subject to specificity of the problem "fair "(in a sense) way of aggregation of this information is built in the "consensus " ranking of all companies. The procedure is applied to real data on ratings assigned by CRA to Russian banks in the period of 2010-2016 in national scales. The calculations show that the obtained consensus ranking is highly stable and can be successfully used as an explanatory (scoring) variable to assess the credit risk of companies. The proposed method is rather common, and it allows to aggregate rankings of arbitrary nature (CRA ratings, internal expert estimates, market indicators), as well as incomplete rankings (for example, when companies are rated only part of the CRA or do not have market debt instruments).

KEYWORDS: credit ratings, expert information, consensus ranking, Kemeny median, lexicographic optimization

JEL Classification: C43, G24 **Received: 01.09.2017 / Published: 31.09.2017**

For correspondence: Buzdalin A.V. (Aleksey.buzdalin@gmail.com)

CITATION:

Buzdalin A.V., Zanochkin A.Yu., Kurbangaleev M.Z., Smirnov S.N. (2017) Agregatsiya kreditnyh reytingov kak zadacha postroeniya konsensusa v sisteme ekspertnyh otsenok [Aggregation of credit ratings as a task of building consensus in the system of expert assessments] Global markets and Financial Engineering. 4(3). – 181-207. doi: 10.18334/grfi.4.3.38830

Обычно при построении сводного показателя параметры стремятся калибровать к историческим данным о дефолтах. Однако нередко возникают случаи, когда стандартные методы (особенно параметрические) не могут дать надежные результаты, прежде всего, по причине малого объема статистических данных: как самих наблюдений, так и событий дефолтов. Так, например, результаты работы коллектива под руководством Смирнова С.Н. «Сравнение качества национальных рейтингов российских банков» (см. [3] *(Afonina et al., 2010)*) показывают, что объем и характер статистики дефолтов на российском рынке без дополнительных предположений не позволяют сделать надежных выводов о разделяющей способности рейтингов, присваиваемых российскими кредитными рейтинговыми агентствами (КРА). Анализ современных данных, публикуемых КРА, показывает, что с тех пор ситуация качественно не изменилась.

Тем не менее агрегирование показателей кредитного качества имеет смысл даже при отсутствии достаточного объема наблюдений. Учитывая, что индивидуальные показатели каждый по своему «зашумлены», агрегированный показатель позволяет добиться большей устойчивости по сравнению с индивидуальными, а значит, будет более надежным для принятия решений.

Другое приложение – построение бенчмарка для валидации рейтинговой системы в условиях низкодефолтного портфеля. В частности, Базельский комитет, среди прочего, рекомендует бенчмаркинг внутренних рейтингов к внешним или внутренним оценкам (см. [4]). Для получения надежного результата агрегирования необходимо принять ряд дополнительных разумных предположений, естественных с экономической точки зрения, и провести предварительную обработку данных с помощью непараметрического инструментария, не прибегая при этом к статистике дефолтов.

В данной статье предлагается подход к агрегации информации об относительном кредитном качестве на основе понятия «консенсуса», опираясь на методы анализа экспертной информации и теории коллективного выбора. В рамках такого подхода агрегированный показатель строится как простое ранжирование компаний, которое является «наиболее непротиворечивым» по отношению к ранжированиям, порождён-

ОБ АВТОРЕ:

Буздалин Алексей Владимирович, заместитель директора Интерфакс-ЦЭА, кандидат экономических наук (Aleksey.buzdalin@gmail.com)
Заночкин Андрей Юрьевич, аналитик (andyzanochkin@gmail.com)
Курбангалеев Марат Зуфарович, младший научный сотрудник (mkurbangaleev@gmail.com)
Смирнов Сергей Николаевич, доцент кафедры системного анализа факультета ВМК МГУ, директор лаборатории по финансовой инженерии и риск-менеджменту НИУ ВШЭ, доцент, кандидат физико-математических наук (s.n.smirnov@gmail.com)

ЦИТИРОВАТЬ СТАТЬЮ:

Буздалин А.В., Заночкин А.Ю., Курбангалеев М.З., Смирнов С.Н. Агрегация кредитных рейтингов как задача построения консенсуса в системе экспертных оценок // Глобальные рынки и финансовый инжиниринг. – 2017. – Том 4. – № 3. – doi: 10.18334/grfi.4.3.38830

ным индивидуальными показателями. При этом противоречием признается ситуация, когда показатели ранжируют пару компаний строго противоположно.

Несмотря на то, что математический инструментарий является достаточно хорошо разработанным (см., например, [5] *(Litvak, 1982)*), в области финансов он применялся редко. Кроме того, данные о кредитных оценках имеют определенную специфику, которая требует адаптации аппарата анализа экспертной информации к задаче оценки кредитного риска. Отдельные рассуждения об ограниченной применимости классического подхода к определению «консенсуса» в задаче агрегирования кредитных рейтингов приводятся в [6] *(Lehmann, Tillich, 2016)*. Предлагаемый подход является новым и представляется достаточно общим, поскольку позволяет включать в качестве индивидуального показателя ограниченную статистику дефолтов, приписывая ей также некоторый вес, или же рыночные данные, позволяющие судить о «мнении рынка» о кредитном риске, например спрэды доходностей облигаций или CDS[4].

Имеются подходы к агрегированию рейтингов, основанные на некоторых специфических моделях. В качестве примера такого подхода можно привести работу [7] *(Grün et al., 2013)*. Авторы предположили, что динамика объективного («реального») кредитного качества объекта представлена ненаблюдаемым случайным процессом с определенными свойствами, интерпретируя этот процесс как некоторый скоринг. Значения скоринговой переменной в конкретный момент времени определяют ранжирование объектов по уровню кредитного риска. Процессы для различных объектов имеют как общую (системную) составляющую, так и индивидуальную для каждого объекта. В каждый момент времени агентства оценивают значения процессов, но с ошибкой, обладающей определенными статистическими свойствами, после чего присваивают рейтинг в зависимости от диапазона значений, в котором оказалась оценка их объективного скоринга. В рамках предлагаемой авторами этой работы модели объективный скоринг можно оценить статистическими методами по наблюдениям за присвоенными рейтингами.

В отличие от [7] *(Grün et al., 2013)* предлагаемый в данной статье подход является нормативным и не опирается на выбор модели. Подход, схожий по смыслу с предлагаемым в данной статье, был представлен в работе [8] *(Eisl et al., 2013)*. В этой статье авторы разрабатывают непараметрический метод для установления соотношения между рейтинговыми шкалами пар рейтинговых агентств. Для каждой рейтинговой категории шкалы каждого рейтингового агентства соотношение шкал отражает степень, с которой категория соотносится с любой рейтинговой категорией другого рейтингового агентства, и таким образом позволяет получить «ре-мэппинг» для рейтинговых категорий двух рейтинговых агентств. Их метод основан полностью на порядковом сопоставлении рейтингов с целью избежать ошибочных оценок вероятностей дефолта или спорных предположений относительно использования рейтин-

[4] Credit Default Swap (производный финансовый инструмент).

гов в качестве четкого показателя кредитного риска[5]. В данном подходе оцениваются соотношения между рейтинговыми категориями для пар рейтинговых агентств, и, таким образом, информация берется из ранжирования. Авторы приходят к выводу о том, что шкалы «Большой тройки»[6] рейтинговых агентств не являются идентичными, находят соотношения между рейтинговыми категориями и показывают значимость учета ранжирования в рейтинговых шкалах при построении бенчмарков. Этот вывод вполне разумен, учитывая разницу в методологиях различных КРА.

При описании подхода, представленного в настоящей статье, будет использован пример агрегирования кредитных рейтингов, присваиваемых КРА. Однако подход является весьма общим, поскольку позволяет агрегировать рейтинговые оценки или ранжирования по риску из произвольных источников (рейтинговых агентств, экспертов, рыночных показателей) и может применяться в условиях неполноты данных, когда компании участвуют не во всех индивидуальных ранжированиях (например, не рейтингуются некоторыми агентствами). В статье приводится обсуждение специфических свойства рейтингов как экспертной информации, описывается постановка задачи агрегации рейтингов как задачи поиска консенсуса в экспертной системе, а также обсуждение свойств решения. При этом внимание концентрируется на национальных рейтинговых шкалах. Способ нахождения численного решения также обсуждается, но не является основной темой статьи.

1. Кредитные рейтинги как экспертные оценки

Кредитные рейтинги отражают мнение агентства об относительной способности объектов отвечать по своим обязательствам. Без возможности количественной оценки качества шкалы каждого рейтингового агентства, прежде всего, выраженного в вероятностях дефолтов в каждой рейтинговой категории, кредитные рейтинги стоит интерпретировать лишь как ранжирование объектов и обязательств по степени кредитного риска, а саму шкалу воспринимать как порядковую («больше», «меньше», «равно»). Более того, ведущие КРА[7] придерживаются именно такой интерпретации (см. [9] *(Kiff, 2012)*).

Несмотря на то, что внешние рейтинги напоминают часто встречающиеся экспертные оценки (например, судейские решения в спорте или рейтинги успеваемости студентов), кредитные оценки, рассматриваемые как экспертная информация, имеют

---

[5] На самом деле при присвоении рейтингов КРА так или иначе приходится опираться в том числе на субъективные оценки тех или иных характеристик объектов.

[6] Неформальное, но часто используемое обозначение наиболее крупных международных кредитных рейтинговых агентств: Fitch, Moody's и Standard&Poors.

[7] Например, в документе S&P Global Rating Definition 2016 года в пункте 61 написано: "S&P Global Ratings' national scale credit ratings are an opinion of an obligor's creditworthiness (issuer, corporate, or counterparty credit rating) or overall capacity to meet specific financial obligations (issue credit rating), relative to other issuers and issues in a given country or region."

ряд специфических особенностей, которые определяют как общую структуру подхода к задаче агрегирования такой информации, так и возможные пути ее решения. Эти особенности выражаются в том, что кредитные оценки оказываются основанными на более разнородной информации по сравнению с указанными примерами. Причины разнородности кроются в следующих характеристиках рейтингов.

1. Определения дефолтов. Все рейтинговые агентства, так или иначе, оценивают возможность дефолта клиента, но под дефолтом могут понимать разные события. Например, некоторые агентства могут рассматривать санацию банков в качестве дефолта, а другие нет. Кроме того, некоторые агентства опираются только на фактологию события, в то время как другие используют собственные суждения о вероятности наступления события. В итоге рейтинги кредитных рейтинговых агентств, оценивающих одну компанию, могут иметь различное смысловое содержание, хотя все они, так или иначе, отражают мнение об относительном кредитном риске.

2. Горизонты кредитного риска. Рейтинговые агентства могут оценивать кредитный риск на различных горизонтах и в разных отношениях к экономическому циклу. Эти аспекты определяют динамические свойства рейтинговых оценок, которые в международной литературе получили название «рейтинговой философии». В международной практике принято выделять два основных типа рейтинговой философии: through-the-cycle (TTC) и point-in-time (PIT)[8]. Таким образом, даже если рейтинговые оценки опираются на одно и то же определение дефолта, они могут быть неоднородны в отношении горизонта риска и устойчивости к фазе экономического цикла.

3. Шкалы. Обычно в экспертных системах эксперты пользуются единообразными описаниями и шкалами для выставления оценок или ранжирования объектов. Однако в совокупности в отношении кредитных рейтингов это свойство типично не выполняется. Рейтинги отличаются не только количеством категорий, но и точкой отсчета. Так, международные шкалы измеряют кредитное качество относительно «абсолютного нуля», коим выступает «безрисковый» или глобально наиболее надежный заемщик. Национальные шкалы же в качестве точки отсчета используют наиболее надежного заемщика в пределах определенной юрисдикции, как правило, государство. Поскольку кредитное качество государства, а также общеэкономические факторы, определяющие положение национальных компаний в глобальной выборке, могут меняться во времени, стационарность соответствия международной и национальных шкал, строго говоря, не гарантирована.

4. Момент актуализации и синхронизация. Процесс рейтингования является динамическим процессом и предполагает собой оценку объекта в разные моменты времени, однако рейтинговые агентства присваивают рейтинги неодновременно.

[8] Подробнее о рейтинговой философии см., например, [2]. Названия типов рейтинговой философии приведены на английском языке, поскольку устойчивой терминологии на русском языке не сложилось. Термины можно было бы условно перевести следующим образом: TTC как «за цикл», PIT как «текущий».

Обычно рейтинги актуализируются (подтверждаются или пересматриваются) по мере поступления актуальной информации о заемщике (прежде всего, финансовой отчетности), а также по мере развития общего состояния экономики, но реакции агентств на новую информацию несинхронны и имеют различную частоту. Отчасти этот недостаток нивелируется грубостью рейтинговой шкалы или же использованием рейтинговой философии TTC, поэтому локальные изменения факторов кредитного качества объекта не требуют изменения кредитного рейтинга. В общем случае чем более грубой является шкала агентства и чем длительнее рассматриваемый горизонт риска, тем реже, при прочих равных условиях, агентству требуется актуализировать свои оценки.

5. Выбор рейтингующих агентств. Выбор компанией (например, эмитентом) рейтингующих ее агентств не является постоянным. Компания самостоятельно определяет количество и состав рейтингующих ее агентств, который зависит от целей компании и ее стратегического поведения в достижении этих целей. Для отдельных целей количество и состав рейтингующих агентств могут быть ограничены требованиями со стороны регуляторов или рыночной инфраструктуры. Например, компании, которые получают рейтинги для выпуска еврооблигаций, должны получать несколько рейтингов от международных агентств, в то время как банкам, желающим привлекать средства федерального бюджета или средства пенсионных накоплений, требуется рейтинг или рейтинги в национальной шкале от аккредитованных Банком России агентств. Поскольку цели компаний и условия их достижения могут меняться, компании могут пересматривать состав рейтингующих их агентств: получать дополнительные рейтинги, переходить от одних агентств к другим, отказываться от услуг агентств, вплоть до полного отказа от рейтингования. Таким образом, состав рейтинговых оценок среди различных компаний в различные моменты времени, вообще говоря, будет неоднородным.

Влияние большинства описанных особенностей можно существенно ограничить селекцией данных, выбирая:

– однотипные рейтингуемые компании (например, банки);

– однотипные сравнимые в разные моменты времени шкалы (например, национальные);

– момент фиксации рейтингов объекта в окрестность отчетных дат (например, ежеквартально).

Однако такие характеристики данных, как неоднородность определений дефолта, различия рейтинговых философий риска и неоднородный состав рейтнгующих агентств, вносят систематические искажения в процесс агрегации рейтинговых оценок, что требуется учитывать при адаптации для решения нашей задачи используемых методов анализа экспертной информации.

2. Обработка и предварительный анализ данных о присваиваемых рейтингах

Далее будем предполагать, что селекция данных была проведена описанным выше способом, тогда рейтинги каждого агентства порождают некоторое отношение на множестве рейтингуемых ими объектами. Рассмотрим совокупность S из n рейтингуемых объектов, т.е. $S = \{1, \ldots, n\}$, мнение об уровне кредитного качества которых высказывают m агентств, присваивая им рейтинги (r_i^k, $k = 1, \ldots, m$ и $i = 1, \ldots, n$) в соответствии со своими шкалами. Шкала k-го агентства имеет l_k рейтинговых категорий, нумерацию договоримся вести по возрастанию от высшего (относительного) кредитного качества к низшему.

Отметим, что в случае с кредитными рейтингами под объектом i можно понимать определенную компанию в некоторый (синхронизованный) момент времени. Также без потери общности можно считать, что рейтинги каждого агентства представлены последовательностью подряд идущих целых неотрицательных чисел, где 1 соответствует наивысшему рейтингу, а наибольший номер, равный количеству категорий в шкале данного агентства – наихудшему (дефолтному) рейтингу. Шкалы агентств могут состоять из разного количества категорий. При этом номера рейтингов выступают исключительно именами категорий и задают их порядок в конкретной шкале, поэтому из одинаковых номеров категорий из разных шкал не следует, что входящие в них объекты имеют одинаковое кредитное качество.

Отношение, порождённое рейтингами некоторого агентства, обладает определенными свойствами, описанными далее в разделе 3 статьи. Эти свойства позволяют применить математический аппарат, используемый в анализе и обработке экспертной информации. Одновременно с этим, как уже отмечалось, различия определений дефолтов и горизонтов рисков вносят систематические искажения в соотношения индивидуальных рейтинговых оценок, а неоднородный состав рейтингующих агентств делает соотношение рейтинговых оценок различных агентств частично неопределенными. Поэтому для применения указанного аппарата следует убедиться для конкретных данных в том, что в принципе имеет смысл искать соответствие между оценками.

Для этой цели удобно использовать матрицы числа совместных наблюдений *(табл. 1)*, содержащие количество объектов, рейтингуемых двумя агентствами, а также матрицы ранговых корреляций (например, модифицированный коэффициент τ_x Кендалла, см. подробнее [10] *(Emond, Mason, 2002)*) на множестве совместных наблюдений *(табл. 2)*, которые показывают, насколько рейтинговые оценки согласуются друг с другом. Модифицированный коэффициент τ_x Кендалла рассчитывается на основе матриц отношений Q^k следующего вида:

$$Q^k = \{q_{ij}^k\}_{n \times n}, \; k = 1, \ldots, m;$$

$$q_{ij}^{k}=\begin{cases} 1,\ r_i^k \geq r_j^k \text{ и } i \neq j; \\ -1,\ r_i^k < r_j^k; \\ 0,\ i=j; \end{cases}$$

и задается формулой:

$$\tau_x\left(Q^k, Q^l\right)=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^{k} q_{ij}^{l}}{n(n-1)}.$$

Чем больше совместных наблюдений имеют агентства и чем выше коэффициент ранговой корреляции, тем более согласованы мнения агентств и тем полезнее становится аппарат анализа экспертной информации.

Таблица 1

Пример матрицы объемов совместных наблюдений четырех КРА, основанный на реальных данных

	КРА 1	КРА 2	КРА 3	КРА 4
КРА 1	1890	297	452	455
КРА 2	297	777	199	180
КРА 3	452	199	1074	107
КРА 4	455	180	107	3594

Источник: рассчитано авторами.

Таблица 2

Пример матрицы корреляций Кендалла для рейтингов четырех КРА, основанный на реальных данных

	КРА 1	КРА 2	КРА 3	КРА 4
КРА 1	1,00	0,70	0,76	0,46
КРА 2	0,70	1,00	0,84	0,60
КРА 3	0,76	0,84	1,00	0,51
КРА 4	0,46	0,60	0,51	1,00

Источник: рассчитано авторами.

3. Математическая постановка задачи

Принимая во внимание интерпретацию кредитных рейтингов как простых ранжирований по уровню риска, критерий для сопоставления и агрегации рейтинговых шкал целесообразно сформулировать в терминах простого противоречия ранжирований на основе рейтингов разных агентств. Прочие соображения могут (и даже должны) быть использованы, но их следует воспринимать как вторичные, уточняющие или позволяющие выбрать единственное решение задачи среди множества решений, оптимальных согласно главному критерию. При постановке мы будем опираться на стандартные концепции, теории коллективного выбора и анализа экспертной информации (подробнее, см. [11] *(Brandt et al., 2016)* и [5] *(Litvak, 1982)*). Особое внимание уделим желательным свойствам результата агрегирования и их интерпретации.

Удобно начать с рассмотрения объектов более общих, чем ранжирования, а именно с бинарных отношений на конечном множестве S, т.е. представляющих собой множества вида $R \subseteq S \times S$ (декартово произведение S на S). Метрика Миркина[9] определяется соотношением

$$\rho(R, R') = \nu(R \Delta R'), \tag{1}$$

где ν – считающая мера[10] на $S \times S$, а символ Δ используется для обозначения симметрической разности множеств. Метрика ρ отражает степень несоответствия R и R'. Отметим[11], что эта метрика совпадает с метрикой Кемени:

$$\nu(R \Delta R') = \int \left| I_R - I_{R'} \right| d\nu = \sum_{z \in S \times S} \left| I_R(z) - I_{R'}(z) \right|, \tag{2}$$

откуда, в частности, видно, что ρ – метрика. Кроме того, аддитивность меры ν влечет:

$$\nu(R \Delta R') = \nu(R) + \nu(R') - 2\nu(R \cap R'). \tag{3}$$

Метрика ρ обладает, в частности, следующим свойством аддитивности: если $R' \subseteq R'' \subseteq R'''$, то $\rho(R', R''') = \rho(R', R'') + \rho(R'', R''')$.

Обозначим $\mathcal{R}_S$ класс всех отношений на S. Пусть заданы отношения $R_1, \ldots, R_m$ на $S \times S$ и $\mathcal{R} \subseteq \mathcal{R}_S$ – класс отношений на $S \times S$ (обычно обладающих определенными свойствами). Медианным классом $\mathcal{R}^* \subseteq \mathcal{R}$ назовем множество тех $R^* \in \mathcal{R}$, которые минимизируют по R функционал Кемени:

$$d(R; R\,, \ldots, R_m) = \sum \ (R, R_k), \tag{4}$$

[9] См. например [12] *(Mirkin, Fenner, 2016)*.

[10] Т.е. $\nu(A) = |A|$ – число точек в множестве A.

[11] Здесь I_A – индикатор множества A, т.е. $I_A(x) = 1$, если $x \in A$, и $I_A(x) = 0$, если $x \notin A$.

т.е. таких, что:

$$d(R^*; R_1, \dots, R_m) = \min_{R \in \mathcal{R}} d(R; R_1, \dots, R_m). \tag{5}$$

Задача тривиальна в случае $m = 1$; далее будем предполагать, что $m > 1$. Отношение $R^* \in \mathcal{R}^*$ будем называть медианой Кемени[12] в классе $\mathcal{R}$ для отношений $R_1, \dots, R_m$. Вид функционала (4) гарантирует свойство анонимности (инвариантность относительно перенумерации R_k, $k = 1, \dots, m$), т.е. все отношения R_k равноправны при определении $\mathcal{R}^*$; очевидно также и свойство нейтральности (инвариантность относительно способа нумерации S).

Из свойства вытекает, что:

$$d(R; R_1, \dots, R_m) = m\nu(R) + \sum_{k=1}^{m} \nu(R_k) - 2\sum_{k=1}^{m} \nu(R \cap R_k);$$

однако $R_1, \dots, R_m$ считаются заданными, поэтому нахождение минимизаторов функционала равносильно поиску минимизаторов для функционала:

$$\psi(R) = m\nu(R) - 2\sum_{k=1}^{m} \nu(R \cap R_k). \tag{6}$$

Обозначим:

$$\varkappa(z) = \sum_{k=1}^{m} I_{R_k}(z),$$

т.е. $\varkappa(z)$ – количество отношений, для которых выполняется $z \in R_k$, $k = 1, \dots, m$;

$$R_{\geq t} = \{z \in S \times S : \varkappa(z) \geq t\}.$$

Пусть $\mathcal{R} = \mathcal{R}_S$; тогда задачу оптимизации (5) будем называть безусловной. Рассмотрим эффект, оказываемый на значение функционала ψ, задаваемого (6), посредством добавления или удаления элемента из отношения $R \in \mathcal{R}_s$.

Пусть $z \notin R$, тогда:

$$\begin{aligned}\psi(R \cup \{z\}) &= \nu(R) + m - 2\sum_{k=1}^{m} \left[\nu(R \cap R_k) + \nu(\{z\} \cap R_k)\right] = \\ &= \psi(R) + m - 2\varkappa(z).\end{aligned} \tag{7}$$

[12] В том числе не исключается, что $R^* = \varnothing$; в частности, если R_k, $k = 1, \dots, m$ (где $m > 1$) попарно не пересекаются, то $R^* = \varnothing$ – медиана.

Если же $z \in R$, тогда:

$$\psi(R \setminus \{z\}) = \nu(R) - m - 2\sum_{k=1}^{m}\left[\nu(R \cap R_k) - \nu(\{z\} \cap R_k)\right] =$$
$$= \psi(R) - m + 2\varkappa(z). \tag{8}$$

Тем самым (7) показывает, что, добавляя точки, для которых $\varkappa(z) \geq \frac{m}{2}$, мы не увеличиваем значения функционала ψ, а если $\varkappa(z) > \frac{m}{2}$, то уменьшаем его значение. Аналогично (8) показывает, что, удаляя точки, для которых $\varkappa(z) \leq \frac{m}{2}$, мы не увеличиваем значения функционала ψ, а если $\varkappa(z) < \frac{m}{2}$, то уменьшаем его значение. Отсюда легко сделать вывод, что для случая, когда m – четно, т.е.[13] $\frac{m}{2} = \left[\frac{m}{2}\right]$, медианный класс $\mathcal{R}^*$ состоит из всех отношений R^*, удовлетворяющих включениями:

$$R_{\geq \frac{m}{2}+1} \subseteq R^* \subseteq R_{\geq \frac{m}{2}}. \tag{9}$$

Из (9) и свойства аддитивности метрики ρ получаем также:

$$\rho(R_{\geq \frac{m}{2}+1}, R^*) + \rho(R^*, R_{\geq \frac{m}{2}}) = \rho(R_{\geq \frac{m}{2}+1}, R_{\geq \frac{m}{2}});$$

причем:

$$\rho(R_{\geq \frac{m}{2}+1}, R_{\geq \frac{m}{2}}) = \nu(R_{\geq \frac{m}{2}} \setminus R_{\geq \frac{m}{2}+1}) = \nu(R_{= \frac{m}{2}}),$$

где обозначено:

$$R_{=t} = \{z \in S \times S : \varkappa(z) = t\}.$$

Для случая, когда m – нечетное, т.е. $\frac{m}{2} = \left[\frac{m}{2}\right] + \frac{1}{2}$, медианный класс $\mathcal{R}^*$ состоит из единственного отношения:

$$R^* = R_{\geq \left[\frac{m}{2}\right]+1}. \tag{10}$$

Соотношения (9) и (10)[14] позволяют говорить о выполнении для R^* «принципа большинства» Кондорсе. Эти же соотношения, в определенной степени, оправдывают название «медиана» для R^*.

---

[13] $[a]$ обозначает целую часть вещественного числа a.
[14] Отметим, что эти соотношения, в частности, влекут справедливость принципа Парето для $R^* \in \mathcal{R}^*$: $R_{\geq m} \subseteq R^* \subseteq R_{\geq 1}$.

Для задачи условной оптимизации (5), т.е. когда $\mathcal{R} \neq \mathcal{R}_S$, соотношения (9) и (10), однако, не обязаны выполняться, поскольку добавление или же исключение одной точки $z \in S \times S$ из $R^* \in \mathcal{R}$ может приводить к потере свойств, задающих $\mathcal{R}$. Поэтому задача условной оптимизации сложнее задачи безусловной оптимизации.

Отметим, что как на русском, так и на английском языке имеются расхождения в терминологии, касающейся свойств порядков. Мы будем, для определенности, придерживаться терминологии, принятой в книге [5] *(Litvak, 1982)*.

В задаче, касающейся рейтингования компаний кредитными рейтинговыми агентствами, имеется определенная специфика, которую мы отразим в следующей формализации. Пусть S – множество компаний, имеющих рейтинги от данных m агентств. Рейтинги k-го агентства, $k = 1,\ldots,m$ порождают частичный порядок R_k (на S) специального вида. Обозначим $S_k \subseteq S$ – множество компаний, имеющих рейтинги от k-го агентства, $k = 1,\ldots,m$. Тогда для $x \in S \setminus S_k$ и любого $y \in S$, отличного от x, будет выполняться $(x, y) \notin R_k$ и $(y, x) \notin R_k$, т.е. такая компания x несравнима с любой другой компанией. Если же рассмотреть сужение R_k на S_k, $k = 1,\ldots,m$, задаваемое как множество пар $(x, y) \in S_k \times S_k$, для которого рейтинг компании x не меньше (по шкале с возрастанием значения рейтинга при убывании кредитного качества) рейтинга компании y, то R_k на S_k представляет собой (нестрогое) ранжирование, т.е. рефлексивное[15], транзитивное[16] и связное[17] отношение. Иными словами, в нашей задаче отношение R_k задается в виде:

$$R_k = \{(i, j) \in S_k \times S_k : r_i^k \geq r_j^k\}, \;\; k = 1,\ldots,m,$$

а функция полезности $u_k(x) = r_x^k$, $x \in S_k$.

Постановку оптимизационной задачи (5) можно, в принципе, сделать более гибкой, отказавшись от свойства анонимности и введя в (4) веса ϕ_k, $k = 1,\ldots,m$, где ϕ_k задает вес расстояния между полным ранжированием и ранжированием k-ого агентства, $\phi_k > 0$ и $\sum_{k=1}^{m} \phi_k = 1$:

$$d'(R; R_1,\ldots,R_m) = \sum_{k=1}^{m} \phi_k \rho(R, R_k). \tag{4'}$$

[15] Рефлексивность означает, что для любого $x \in S$ $(x, x) \in R$.

[16] Транзитивность R означает, что для x, y и z из S $(x, y) \in R$ и $(y, z) \in R$ влечет $(x, z) \in R$.

[17] Связность означает, что для любых x и y из S верно хотя бы одно из отношений $(x, y) \in R$ и $(y, x) \in R$.

Если оснований[18] доверять одному агентству больше, чем другому, нет, то следует выбрать $\phi_k = 1/m$ для любого $k = 1,\ldots,m$, что соответствует свойству анонимности[19]. Для задания отношения $R \in \mathcal{R}_S$ удобно пользоваться ориентированными графами с дугой из вершины $x \in S$ в вершину $y \in S$, если $(x, y) \in R$, либо же матрицей P инцидентности этого графа:

$$p_{ij} = \begin{cases} 1, \text{ если } (i, j) \in R; \\ 0, \text{ если } (i, j) \notin R. \end{cases}$$

Другими словами, $p_{ij} = I_R\big((i, j)\big)$, так что (2) можно переписать в виде:

$$\rho(R, R') = \sum_{i=1}^{n}\sum_{j=1}^{n} |p_{ij} - p'_{ij}|, \tag{2'}$$

где P и P' – матрицы инцидентности для R и R' соответственно R'.

Наблюдения (эмпирические данные) представляют собой для $x \in S$ наборы вида $U(x) = (u_1(x),\ldots,u_m(x))$, условно полагая $u_k(x) = NR$ в случае, если $x \notin S_k$. Разумеется, для одинаковых наборов вида $U(x) = W$, где $W = (w_1,\ldots,w_m)$, нет оснований различать компании по кредитному качеству, так что имеет смысл агрегировать наблюдаемую информацию по группам компаний вида:

$$U^{-1}(W) = \{x \in S : U(x) = W\},$$

т.е. объединять компании в группы с одинаковыми комбинациями рейтингов W. Оптимальное решение R^* задачи (5), тем самым, разумно искать среди ранжирований, порожденных функцией полезности $u^* : S \mapsto \mathbb{R}$ вида:

$$u^*(x) = v(U(x)) \tag{11}$$

для некоторой функции $v : U(S) \mapsto \mathbb{R}$.

Заметим, что информация, касающаяся соотношения кредитного качества компаний x и y тем надежнее (и тем самым ценнее), чем больше рейтингов имеют эти компании от различных КРА. В частности, единственный рейтинг компании может оказаться результатом практики «rating shopping», при которой получатель рейтинга

[18] Основанием для присваивания мнению одного агентства большего веса по сравнению с другими может быть, например, существенно большая разделяющая способность его рейтинговой системы. Для непосредственной оценки разделяющей способности рейтинговой системы необходимо как минимум располагать статистикой дефолтов среди компаний, отрейтингованных по данной системе, или иметь информацию для косвенной оценки. Если для априорной приоритезации мнений агентств нет возможности или необходимости, естественно считать мнения агентств равносильными.

[19] Мультипликативная константы в (4') не изменяет минимизаторы по сравнению с [4].

запрашивает рейтинги различных агентств, но публикует только наиболее благоприятный для него.

Такая практика является более привлекательной и эффективной в тех случаях, когда компания сама оплачивает услуги агентства по присвоению рейтинга и когда для целей компании не требуется несколько рейтингов. Поэтому, с целью повышения надежности информации, разумно ограничиться рассмотрением только тех компаний, которые имеют как минимум два рейтинга различных КРА. К тому же, хотя (9) и не обязано, вообще говоря, выполняться, но это интуитивно подсказывает, что наиболее полезную информацию для конструкции R^* дают пары (x,y), для которых $\varkappa((x,y))$ достаточно велико, порядка $\frac{m}{2}$ или более, что характерно для тех x и y, у которых имеются рейтинги от различных КРА (множественные рейтинги)[20]. Учитывая также, что единичные рейтинги не несут информации о расхождении мнений различных агентств, имеет смысл постановка задачи, для которой в S входят только компании, имеющие множественные рейтинги, т.е. не менее двух рейтингов различных КРА.

Что касается класса $\mathcal{R}$ отношений, фигурирующего в задаче условной минимизации (5), его естественно считать некоторым классом (нестрогих) ранжирований. Возможны несколько постановок, представляющие интерес в контексте различной интерпретации рейтинговой информации.

В первом варианте постановки, если для k-го агентства имеет место равенство $u_k(x)=u_k(y)$, то считаем, что x и y эквивалентны (отношение эквивалентности при этом задается на S_k), и это учитывается при расчете метрики $\rho(R,R_k)$.

Во втором варианте постановки модифицируем R_k, исключая те пары $(x,y)\in S\times S$, для которых $u_k(x)=u_k(y)$, получая тем самым новое отношение $R'_k\subseteq R_k$, после чего будем искать решение оптимизационной задачи (5) для $R'_1,\ldots,R'_m$ вместо $R_1,\ldots,R_m$. Второй вариант можно интерпретировать таким образом: для компаний с одинаковым рейтингом k-го агентства методика этого агентства слишком груба, чтобы распознать, у какой из компаний кредитное качество выше. Второй вариант может оказаться предпочтительней, исходя из нижеследующих соображений.

Пусть отношения R' и R'' являются (нестрогими) ранжированиями на $S'\subseteq S$ и $S''\subseteq S$ соответственно. Если для пары $(i,j)\in S\times S$ выполняется $i\in S\setminus S'$, то будем говорить, что пара (i,j) R'-несравнима, а если $j\in S\setminus S''$, то будем говорить, что пара (i,j) R''-несравнима.

Обозначим $B=B(R',R'')$ матрицу с элементами:

---

[20] Следует отметить, что на практике полезность множественных рейтингов зависит от специфики и характера данных. Во-первых, наблюдения о множественных рейтингах преимущественно состоят из компаний, рейтингуемых двумя агентствами (например, в данных, используемых в данной статье, такие наблюдения составляют 74 %). Во-вторых, если рейтинги сильно отличаются по своей природе (см. раздел 1 данной статьи), то чем больше рейтингов присвоено компаниям, тем вероятнее получить существенный уровень разногласий мнений агентств.

$$B_{ij}(R',R'') = | p'_{ij} - p''_{ij} | + | p'_{ji} - p''_{ji} |,$$

где матрицы P' и P'' с элементами p'_{ij} и p''_{ij}, $i \in \{1,\ldots,n\}$, $j \in \{1,\ldots,n\}$ являются матрицами инцидентности для R' и R''. При этом, в силу свойств R' и R'', найдутся функции полезности u' и u'', такие, что $(i,j) \in R'$ равносильно $u'(i) \ge u'(j)$ и $(i,j) \in R''$ равносильно $u''(i) \ge u''(j)$. Не ограничивая общности, можно считать, что $u'(S) = [1,l']$, $l' \le \ (S') \le n$ и $u''(S) = [1,l'']$, $l'' \le \nu(S'') \le n$, где $[\alpha, b]$ обозначает интервал всех целых чисел между α и b (включительно). Обозначим для таких функций полезности[21] $r'_i = u'(i)$ и $r''_i = u''(i)$. В этом случае:

$$\rho(R',R'') = \sum_{i<j} B_{ij}(R',R'') = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}(R',R'')\cdot$$

Нетрудно проверить, что:

$B_{ij}(R',R'') = 2$ в случаях, если:

1) $r'_i > r'_j$ и $r''_i < r''_j$,
2) $r'_i < r'_j$ и $r''_i > r''_j$,
3) $r'_i = r'_j$ и пара (i, j) R''-несравнима,
4) пара (i, j) R'-несравнима и $r''_i = r''_j$,

$B_{ij}(R',R'') = 1$ в случаях, если:

5) пара (i, j) R'-несравнима и $r''_i > r''_j$,
6) пара (i, j) R'-несравнима и $r''_i < r''_j$,
7) $r'_i = r'_j$ и $r''_i < r''_j$,
8) $r'_i < r'_j$ и $r''_i = r''_j$,
9) $r'_i > r'_j$ и пара (i, j) R''-несравнима,
10) $r'_i < r'_j$ и пара (i, j) R''-несравнима,

$B_{ij}(R',R'') = 0$ в случаях, если:

11) пара (i, j) R'-несравнима и пара (i, j) R''-несравнима,
12) $r'_i = r'_j$ и $r''_i = r''_j$,
13) $r'_i > r'_j$ и $r''_i > r''_j$,
14) $r'_i < r'_j$ и $r''_i < r''_j$.

Отметим, что если R' является (нестрогим) ранжированием на всем S (т.е. $S' = S$), тогда количество вариантов, определяющих значение $B_{ij}(R',R'')$, сокращается с 14 до 10, поскольку варианты 4), 5), 6) и 11) в этом случае невозможны.

По своему смыслу матрица $B(R',R'')$ отражает штраф за несоответствие отношений R' и R''. В этом плане случаи 3) и 4) выглядят наиболее неестественным образом. Вторая постановка решает эту проблему, поскольку тогда случай $r'_i = r'_j$ равносилен R'-несравнимости пары (i, j), а случай $r''_i = r''_j$ равносилен R''-несравнимости пары $(i,$

[21] По сути дела, речь идет о «рейтинговых шкалах» для R' и R''.

j). Однако вторая постановка все же не решает проблему для случаев 5), 6), 9) и 10), которые также выглядят неестественно. Кроме того, учитывая предложенную интерпретацию, случаи 7) и 8) во второй постановке также не выглядят естественно.

Исходя из обозначенных выше проблем с интерпретацией, можно пойти дальше в модификации постановки задачи[22] и ввести вместо метрики ρ невязку σ, полагая:

$$\sigma(R', R'') = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij}(R', R''), \tag{12}$$

где:

$$C_{ij}(R', R'') = \begin{cases} 1, \text{ если } r_i' > r_j' \text{ и } r_i'' < r_j'' \text{ или } r_i' < r_j' \text{ и } r_i'' > r_j''; \\ 0, \text{ в противном случае.} \end{cases} \tag{13}$$

Тем самым $C_{ij}(R', R'') = 1$ только в случаях a) и b), а в остальных восьми случаях $C_{ij}(R', R'') = 0$. Отметим, однако, что невязка σ, вообще говоря, не является метрикой.

Пусть $\mathcal{R}'$ - класс всех (нестрогих) ранжирований, заданных на всем S, порожденных функциями полезности u^* вида с инъективной[23] функцией v. Задача минимизации по $R \in \mathcal{R}'$ нового функционала:

$$d'(R; R_1, \ldots, R_m) = \sum_{k=1}^{m} \sigma(R, R_k), \tag{14}$$

где σ задается соотношениями (12) и (13), является вполне осмысленной постановкой в нашей задаче, с учетом ее специфики. Минимизатор R^* в (14) будем называть квазимедианой.

Поскольку невязка σ грубее по сравнению с метрикой ρ – в меньшей степени реагирует на различия в бинарных отношениях, – то множество квазимедиан (т.е. решений задач (14), по крайней мере, для второй постановки, шире, чем множество медиан Кемени (т.е. решений задачи (5)). Таким образом, задачи (5) и (14) приводят к неединственности искомого ранжирования.

Несколько факторов влияют на количество ранжирований – медиан или квазимедиан – и, следовательно, вычислительную сложность нахождения их всех. При прочих равных условиях множество решений тем шире, чем менее согласованны мнения рейтинговых агентств между собой, точнее, имеется достаточно обширное количество

[22] Для этой модификации безразлично, какой вариант постановки (в отношении включения или исключения пар с одинаковой функцией полезности) выбирается – первый или второй, поэтому далее будем предполагать для определенности, что речь идет о первом варианте.

[23] Можно сказать, что соответствующее упорядочение классов вида $U^{-1}(W)$, где $W \in U(S)$, образует строгое ранжирование, т.е. линейный порядок (кроме рефлексивности, транзитивности и связности, также выполняется свойство антисимметричности, т.е. если $(x, y) \in R$ и $(y, x) \in R$, то $x = y$).

«спорных» случаев. Число ранжирований – медиан и квазимедиан – также увеличивается, когда возникает парадокс Кондорсе, т.е. противоречия по принципу большинства в случае циклов (цикл длины 3 означает: α_1 предпочитается α_2, α_2 предпочитается α_3, а α_3 предпочитается α_1)[24]. Например, в работе [13] *(Muravyov, Marinushkina, 2013)* показано, что шанс столкнуться с парадоксами Кондорсе выше в случаях, когда число экспертов, чьи мнения агрегируются, является нечетным. Наконец, нестрогость порядков среди ранжирований R_k, $k = 1, \ldots, m$ также приводит к росту мощности множества медиан Кемени, поскольку объекты с одинаковым набором рангов можно произвольно менять местами. Нестрогость порядков является наиболее легко преодолимой из перечисленных факторов, поскольку устраняется объединением объектов с одинаковыми рангами в группы, после чего ранжирование строится на уровне групп. Таким образом, медианы Кемени и квазимедианы не следует считать готовым агрегатором индивидуальных рейтингов, но они образуют множество «разумных» кандидатов на такую роль[25].

Чтобы выбрать единственное решение, мы предлагаем новую модификацию постановки задачи – сформулировать дополнительный (второстепенный) критерий и решать лексикографическую задачу, т.е. из всех оптимальных по основному критерию ранжирований – медиан Кемени или квазимедиан – выбрать оптимальную по дополнительному критерию. Решение задачи лексикографической оптимизации договоримся называть «уточненной медианой Кемени», или же, соответственно, «уточненной квазимедианой». При этом уточненную квазимедиану естественно искать в более широком, чем $\mathcal{R}'$, классе $\mathcal{R}''$ всех (нестрогих) ранжирований, заданных на всем S и порожденных функциями полезности u^* вида (11), без требования инъективности функции v. В принципе, в задачу также можно вводить иные ограничения.

В качестве дополнительного критерия разумно выбрать некоторый метризованный критерий с невязкой δ^2, описанной ниже, который позволяет уточнить «степень» противоречивости и отсечь ранжирования, крайне противоречивые по отношению к рейтингам отдельного агентства. В отличие от первого критерия, метризованный критерий порождает не медиану, а среднее по Кемени (но на классе найденных медиан или квазимедиан), и его минимизация представляет собой, в принципе, более простую задачу, чем нахождение медианы или квазимедианы. Однако непосредственное численное решение этой задачи вряд ли возможно для данных большой размерности, поскольку нахождение всех медиан или квазимедиан крайне затруднительно.

Поэтому для приближенного поиска уточненных квазимедиан мы предлагаем решать следующую оптимизационную задачу, с метризованным критерием в качестве

[24] В этом случае говорят, что профиль исходных порядков нетранзитивен.

[25] Например, похожая ситуация возникает, когда находят границу для задачи многокритеариальной Парето оптимизации.

«регуляризирующей»[26] компоненты. Речь идет о минимизации по $R \in \mathcal{R}''$ функционала:

$$d''(R; R_1, \ldots, R_m) = \sum_{k=1}^{m} \phi_k [\sigma(R, R_k) + \lambda \, \delta^2(R, R_k)], \tag{15}$$

где δ^2 – метризованный критерий, описанный ниже (см.); $\lambda > 0$; $\phi_k > 0$, $k = 1, \ldots, m$; $\sum_{k=1}^{m} \phi_k = 1$.

Следуя терминологии книги [5] *(Litvak, 1982)*, метризованным бинарным отношением назовем отношение $R \subseteq S \times S$, снабженное числовой функцией $w : R \mapsto \mathbb{R}$, и будем обозначать $w_{ij} = w((i, j))$. Метризованное транзитивное отношение называется аддитивным, если функция w обладает свойством: если $(i, j) \in R$ и $(j, k) \in R$, то:

$$w_{ik} = w_{ij} + w_{jk}. \tag{16}$$

Если это отношение также является рефлексивным, то $w_{ii} = 0$ для всех $i \in S$.

Рассмотрим в качестве (R, w) аддитивное метризованное нестрогое ранжирование (добавилось предположение связности). Продолжим функцию $w : R \mapsto \mathbb{R}$ на S, полагая для $(i, j) \notin R$:

$$w((i, j)) = w_{ij} = -w_{ji}, \tag{17}$$

(в силу связности R пара $(j, i) \in R$); при этом для эквивалентных i и j соотношение (17) выполняется вследствие аддитивности, с учетом $w_{ii} = 0$, поскольку $w_{ii} = w_{ij} + w_{ji}$. Таким образом, матрица $W = (w_{ij}, i = 1, \ldots, m; j = 1, \ldots, n)$ является антисимметричной.

При выбранном способе продолжения функции w свойство аддитивности (16) будет выполняться для любых i, j и k из S. В этом можно убедиться непосредственно, проверяя восемь вариантов для пар (i, j), (i, k) и (k, j), каждая из которых может принадлежать R или нет. При этом два варианта: $(i, j) \notin R$, $(i, k) \in R$, $(k, j) \in R$ и $(i, j) \in R$, $(i, k) \notin R$, $(k, j) \notin R$ невозможны в силу транзитивности.

Аддитивное метризованное (нестрогое) ранжирование возникает тогда и только тогда, когда найдется функция полезности[27] $u : S \mapsto \mathbb{R}$, такая, что $R = \{(i, j) \in S \times S : u(i) \geq u(j)\}$ и

$$w_{ij} = u(i) - u(j). \tag{18}$$

[26] Принцип регуляризации Тихонова состоит в замене исходной задачи на близкую к ней, дающую корректные решения в пределе. Обычно, впрочем, этот принцип понимают более узко, только для некорректно поставленных обратных задач, чтобы добиться устойчивого численного решения.

[27] Отметим, что для строгого ранжирования функция полезности инъективна: в книге [5] *(Litvak, 1982)* рассматриваются строгие ранжирования.

Действительно, в силу свойств (нестрогого) ранжирования такая функция u существует, а задавая w_{ij} по формуле (18), получаем, очевидно, выполнение свойства (16). Обратно достаточно положить:

$$u(i) = \frac{1}{n}\sum_{k=1}^{n} w_{ik} \,. \tag{19}$$

В силу того, что свойство (16) выполняется для любых i, j и k из S, имеем $w_{ij} = w_{ik} - w_{jk}$ для любого $k = 1,\ldots,n$. Тем самым из (19) получаем:

$$u(i) - u(j) = \frac{1}{n}\sum_{k=1}^{n}(w_{ik} - w_{jk}) = w_{ij} \,.$$

Пусть теперь R' и R'' – (нестрогие) ранжирования на $S' \subseteq S$ и $S'' \subseteq S$ соответственно, $u' : S' \mapsto \mathbb{R}$ и $u'' : S'' \mapsto \mathbb{R}$ соответствующие функции полезности. Зададим w' и w'', полагая, как в , $w'_{ij} = u'(i) - u'(j)$ и $w''_{ij} = u''(i) - u''(j)$. Определим невязку δ^2 ранжирований R' и R'' следующим образом:

$$\delta^2(R', R'') = \sum_{i \in S' \cap S''}\ \sum_{j \in S' \cap S''} (w'_{ij} - w''_{ij})^2 \,. \tag{20}$$

Для целей нашего приложения мы выбираем u' и u'' таким образом, чтобы $l'u'(S') = [1, l']$, $l' \le \nu(S')$ и $l''u''(S'') = [1, l'']$, $l'' \le \nu(S'')$. Отметим, что в формуле (15) ранжирование R определено на всем S, так что суммы по i и j в (20) берутся по множеству S_k, $k = 1,\ldots,m$.

При большой размерности задачи проводить оптимизацию очень затруднительно с вычислительной точки зрения[28] и вряд ли можно надеяться, что удастся найти точное решение даже на мощных суперкомпьютерах. Решение подобной задачи приходится искать эвристическими численными методами; опираясь на современный опыт, был выбран некоторый класс меметических алгоритмов. Помимо проверенной эффективности класс таких алгоритмов может быть приспособлен для параллельных вычислений. Подобный алгоритм был адаптирован под специфику нашей задачи и реализован в среде MATLAB с использованием возможностей параллельных вычислений. Описание алгоритма лежит за пределами данной статьи и является предметом отдельной работы. Здесь мы лишь приведем ссылку на литературу (см. [14] *(Martí, Reinelt, 2011)*), где описаны соответствующие подходы к выбору алгоритмов.

[28] Рассматриваемая задача относится к классу NP-сложных задач (этот результат представляет обобщение частных результатов, полученных различными авторами за продолжительный период времени; общую теорему см. в разделе 4.2.1 в [11] *(Brandt et al., 2016)*).

4. Иллюстрация агрегирования рейтингов на примере рейтингов российских банков в национальных шкалах

Будем называть везде далее решение задачи (15) консенсусным ранжированием. Проиллюстрируем построение консенсусного ранжирования на примере агрегации рейтингов российских банков, имеющих два или более рейтингов в национальных шкалах. Используются квартальные данные о рейтингах, присвоенных семью рейтинговыми агентствами в период[29] с середины 2010 г. по середину 2016 г. Для расчетов была принята конвенция, при которой наблюдение за парой объектов, когда одно агентство присвоило одинаковые рейтинги, а другое – разные, не считается противоречием и не вносит вклад в значение минимизируемого функционала в (15). В расчетах выбирались равные веса ϕ_k, $k = 1, \ldots, 7$.

На рисунке 1 рейтинги одного из агентств отображены относительно результирующего консенсусного ранжирования.

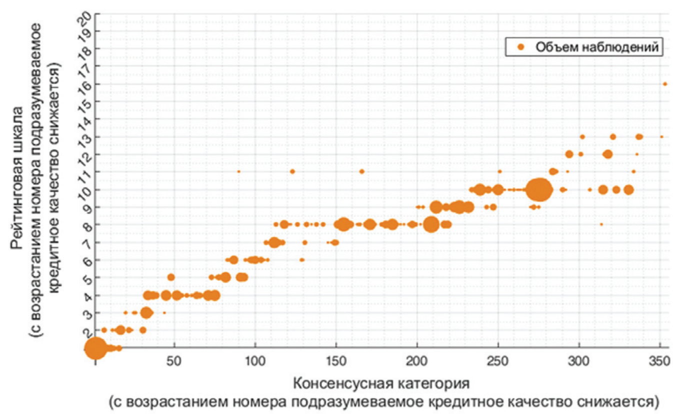

Рисунок 1. Иллюстрация согласованности консенсусного ранжирования и ранжирования на основе отдельных рейтинговых оценок
Источник: составлено авторами.

Как видно из этого рисунка, степень согласованности между консенсусным ранжированием и рейтингами агентства довольно высокая, что выражено в образовании достаточно четкого восходящего ступенчатого характера отображения. Рисунок также

[29] Начало используемого исторического периода соответствует вступлению в силу приказа Минфина России от 4 мая 2010 г. № 37н, который установил порядок аккредитации рейтинговых агентств, тем самым задав определенный уровень рыночной дисциплины в рейтинговой индустрии. Изменение среды функционирования КРА в том числе сказалось в целом на качестве и полноте данных о присваиваемых в России рейтингах. Окончание периода определено ограничениями доступности данных для авторов исследования.

показывает, что консенсусное ранжирование содержит существенно больше категорий по сравнению с исходными шкалами: в представленном примере рейтинговые шкалы состояли из 10 или 20 категорий, в то время как консенсусное ранжирование содержит 355 категорий.

В общем случае количество категорий в консенсусном ранжировании тем больше, чем больше уникальных комбинаций рейтингов встречается в данных, а их порядок зависит от степени противоречивости каждого набора рейтингов остальному массиву данных. Отметим, что если бы ситуации, когда одно агентство одинаково оценивает пару объектов, а другое – строго приоретизирует, считалась бы противоречием, тогда количество категорий в консенсусном ранжировании было бы значительно меньшим[30], консенсус в существенной степени ориентировался бы на рейтинги агентства с наименее подробной шкалой, и рейтинги такого агентства ограничивали бы возможности учета информации из более подробных рейтингов в консенсусное ранжирование. Напротив, если описанная ситуация не считается противоречием, то консенсус способен учесть информацию из более подробных рейтингов и уточнить рейтинги агентств, использующих менее подробные шкалы. Таким образом, в зависимости от нюансов постановки, консенсусное ранжирование способно по-разному агрегировать информацию о кредитном качестве из состава рейтинговых оценок.

Следует отметить, что процедура поиска консенсусного ранжирования приводит к довольно устойчивому решению, что демонстрируют рисунки 2 и 3.

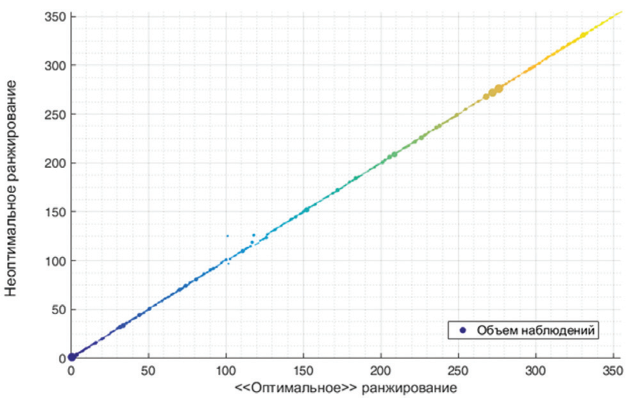

Рисунок 2. Иллюстрация устойчивости алгоритма поиска консенсусного ранжирования: неоптимальное ранжирование против «оптимального»
Источник: составлено авторами.

[30] На наших данных получилось 47, а не 355 категорий кредитного качества в консенсусном ранжировании.

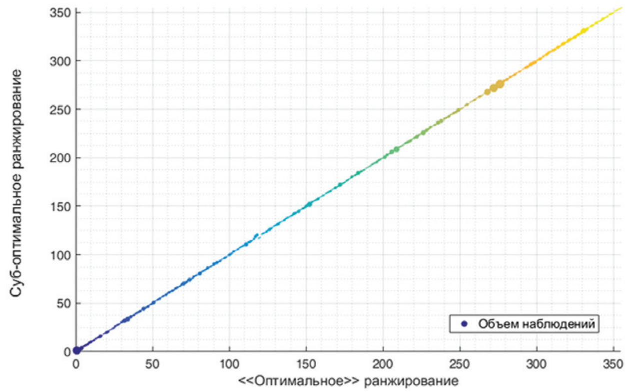

Рисунок 3. Иллюстрация устойчивости алгоритма поиска консенсусного ранжирования: субоптимальное ранжирование против «оптимального»
Источник: составлено авторами.

На рисунке 2 представлено взаимное положение ранжирования выбранного в качестве «оптимального» агрегатора и неоптимального решения, которое в терминах критерия находится сильно далеко от агрегатора. Из рисунка видно, что неоптимальное и «оптимальное»[31] ранжирования имеют локальные противоречия в районе 100–120 категорий, однако характеризуются очень высоким уровнем ранговой корреляции (99 % по модифицированному показателю Кендалла). На рисунке 3 показана аналогичная картинка, но для субоптимального ранжирования, которое по критерию находится близко от оптимального. В этом случае ранжирования практически совпадают, а уровень модифицированной ранговой корреляции Кендалла близок к 100 %.

Консенсусное ранжирование может быть интерпретировано как универсальная объясняющая переменная на множестве рейтингуемых объектов, независимо от состава присваиваемых рейтингов. В этом случае естественно задаться вопросом: насколько хорош полученный агрегатор в качестве объясняющей переменной дефолтов. Для ответа на этот вопрос построим CAP-кривую[32] для консенсусного ранжирования на множестве всех рейтингуемых объектов *(рис. 4)*. По критерию[33] «показатель AR» полученное ранжирование имеет разделяющую способность, сопоставимую с разделяющей способностью рейтингов, его составляющих, что позволяет использовать полученное ранжирование для целей валидации внутренних рейтинговых систем банков в условиях низкой дефолтности портфеля.

[31] Под «оптимальным» ранжированием понимается приближенное решение задачи , полученное с помощью выбранного численного метода (алгоритма) при заданном уровне допустимой ошибки.
[32] См. определение в работе [1] *(Tasche, 2010)*.
[33] См. определение этого и других критериев в работе [1] *(Tasche, 2010)*.

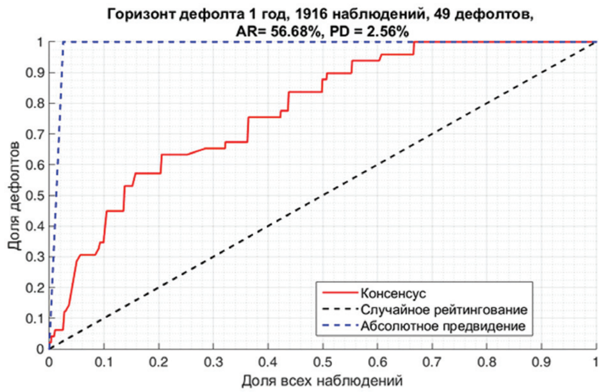

Рисунок 4. CAP-кривая для консенсусного ранжирования
Источник: составлено авторами.

Заключение

В статье представлен подход к агрегации ранжирований объектов по уровню кредитного риска произвольной природы в единое ранжирование. В качестве иллюстрации для описания подхода используется задача агрегирования кредитных рейтингов и используются интерпретации кредитных рейтингов как (нестрогих) ранжирований объектов по уровню их кредитного риска. При такой интерпретации можно построить такое ранжирование объектов рейтингования, которое представляло бы собой «агрегированное мнение» рейтинговых агентств.

В основу подхода положен инструментарий теории коллективной выборки и анализа экспертной информации, а также принятые в них естественные требования к свойствам агрегирования, которые можно интерпретировать как свойства «справедливого», агрегированного ранжированием по отношению к отдельным агентствам. Следуя терминологии теории коллективного выбора, «справедливое» агрегированное ранжирование мы называем «консенсусным».

Класс ранжирований с желательными свойствами дает совокупность медиан Кемени, которая представляет собой множество решений задачи минимизации суммы расстояний Кемени между искомым ранжированием и ранжированиями объектов рейтингования, отвечающими рейтингам отдельных агентств. Специфика задачи, однако, приводит к целесообразности модификации минимизируемого функционала; его минимизаторы мы называем квазимедианами. Обычно медиана Кемени, а также квазимедиана, неединственна, поэтому для выбора конкретного решения в работе предложен подход к его уточнению с помощью лексикографической оптимизации, задавая дополнительный (второстепенный) критерий оптимизации. В качестве такого критерия выбирается аддитивное метризованное расстояние между ранжированиями или же модификация метризованного расстояния. Таким образом, итоговое решение является наилучшим с точки зрения дополнительного критерия среди медиан Кемени или квазимедиан.

Поскольку по своей природе задача нахождения медианы Кемени обладает высокой вычислительной сложностью, в работе предложена оригинальная постановка для численного решения задачи, с дополнительным критерием в качестве «регуляризирующей» компоненты, а также разработан алгоритм, позволяющий находить «уточненную медиану Кемени» приближенно с приемлемой точностью за разумное время счета на компьютере.

Работоспособность подхода была апробирована на рейтингах, присвоенных российским банкам семью кредитными агентствами в национальной шкале в период с середины 2010 г. по середину 2016 г.

Результаты показывают, что результирующий агрегатор в высокой степени согласован с исходными рейтингами, является устойчивым и имеет разделяющую способность на множестве всех рейтингуемых объектов, сопоставимую с разделяющей способностью рейтингов отдельных агентств.

Подход может быть использован для решения широкого круга задач, например, с целью построения бенчмарка для валидации низкодефолтного портфеля.

ИСТОЧНИКИ:

1. Tasche D. Estimating discriminatory power and PD curves when the number of defaults is small. ArXiv:0905.3928v2. [Электронный ресурс]. URL: http://arxiv.org/abs/0905.3928.
2. Studies on the Validation of Internal Rating Systems. Working Papers of the Basel Committee on Banking Supervision. [Электронный ресурс]. URL: http://www.bis.org/publ/bcbs_wp14.pdf.
3. Афонина С.Г., Богатырева Е А., Косьяненко А.В., Лапшин В.А., Науменко В.В., Смирнов С.Н. Сопоставление качества рейтингов российских банков: Препринт WP16/2010/03. Высшая школа экономики. [Электронный ресурс]. URL: https://fermlab.hse.ru/data/2010/07/28/1218992134/WP16_2010_03-ff.pdf.
4. Validation of low-default portfolios in the Basel II Framework. Bcbs. [Электронный ресурс]. URL: https://www.bis.org/publ/bcbs_nl6.pdf.
5. Литвак Б.Г. Экспертная информация: Методы получения и анализа. – М.: Радио и Связь, 1982. – 184 с.
6. Lehmann C., Tillich D. Consensus Information and Consensus Rating // Operations Research Proceedings 2014: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), RWTH Aachen University. Germany, 2016. – p. 357-362.
7. Grün B., Hofmarcher P., Hornik K., Leitner C., Pichler S. Deriving consensus ratings of the big three rating agencies // Journal of Credit Risk. – 2013. – № 1. – p. 75-98.

8. Eisl A., Elendner H.W., Lingo M. Re-Mapping Credit Ratings. Ssrn.com. [Электронный ресурс]. URL: https://ssrn.com/abstract=1836877.
9. Kiff J., Nowak S., Schumacher L. Are Rating Agencies Powerful? An Investigation into the Impact and Accuracy of Sovereign Ratings. International Monetary Fund (Working Paper). [Электронный ресурс]. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1997736.
10. Emond E.J., Mason D.W. New Rank Correlation Coefficient with Application to the Consensus Ranking Problem // Journal of Multi-criteria decision analysis. – 2002. – № 9. – p. 17-28.
11. Brandt F., Conitzer V., Endriss U. Handbook of Computational Social Choice. - Cambridge University Press, 2016. – 548 p.
12. Mirkin B., Fenner T.I. Tied Rankings, Ordered Partitions, and Weak Orders: WP7/2016/08// WP7 "Mathematical methods for decision making in economics, business and politics". Higher School of Economics Publ. House. [Электронный ресурс]. URL: https://www.hse.ru/data/2016/12/15/1111586825/WP7_2016_08-f.pdf.
13. Muravyov S.V., Marinushkina I.A. Intransitivity in multiple solutions of Kemeny Ranking Problem // Journal of Physics: Conference Series. – 2013. – № 1.
14. Martí R., Reinelt G. The linear ordering problem: exact and heuristic methods in combinatorial optimization // Springer Science & Business Medi. – 2011. – № 175. – p. 172.

REFERENCES:

Brandt F., Conitzer V., Endriss U. (2016). Handbook of Computational Social Choice. Cambridge University Press.

Eisl A., Elendner H.W., Lingo M. Re-Mapping Credit RatingsSsrn.com. Retrieved from https://ssrn.com/abstract=1836877

Emond E.J., Mason D.W. (2002). New Rank Correlation Coefficient with Application to the Consensus Ranking Problem Journal of Multi-criteria decision analysis. 28 (9). 17-28.

Grün B., Hofmarcher P., Hornik K., Leitner C., Pichler S. (2013). Deriving consensus ratings of the big three rating agencies Journal of Credit Risk. 9 (1). 75-98.

Kiff J., Nowak S., Schumacher L. Are Rating Agencies Powerful? An Investigation into the Impact and Accuracy of Sovereign RatingsInternational Monetary Fund (Working Paper). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1997736

Lehmann C., Tillich D. (2016). Consensus Information and Consensus Rating Operations Research Proceedings 2014. 357-362.

Martí R., Reinelt G. (2011). The linear ordering problem: exact and heuristic methods in combinatorial optimization Springer Science & Business Medi. (175). 172.

Mirkin B., Fenner T.I. Tied Rankings, Ordered Partitions, and Weak Orders: WP7/2016/08// WP7 "Mathematical methods for decision making in economics, business and politics"Higher School of Economics Publ. House. Retrieved from https://www.hse.ru/data/2016/12/15/1111586825/WP7_2016_08-f.pdf

Muravyov S.V., Marinushkina I.A. (2013). Intransitivity in multiple solutions of Kemeny Ranking Problem Journal of Physics: Conference Series. 459 (1).

Studies on the Validation of Internal Rating SystemsWorking Papers of the Basel Committee on Banking Supervision. Retrieved from http://www.bis.org/publ/bcbs_wp14.pdf

Tasche D. Estimating discriminatory power and PD curves when the number of defaults is smallArXiv:0905.3928v2. Retrieved from http://arxiv.org/abs/0905.3928

Validation of low-default portfolios in the Basel II FrameworkBcbs. Retrieved from https://www.bis.org/publ/bcbs_nl6.pdf