

Использование существующих подходов к моделированию ожидаемых кредитных убытков во время кризиса, вызванного пандемией COVID-19, может привести к менее точному расчету резервов ECL, что станет причиной чрезмерной или, наоборот, недостаточной аллокации капитала на покрытие кредитных потерь. По мнению авторов, в ближайшее время на первый план выйдут подходы, основанные на алгоритмах машинного обучения и использовании концепции Auto ML. В статье на реальных примерах из практики банков описаны модели, учитывающие специфику кризисного периода.

Алексей АНТОНОВ, компания «Неофлекс», директор практики «Технологии финансовых рынков и управления рисками», к.э.н.

Роман СОРОКИН, компания «Неофлекс», ведущий специалист практики Data Science

Оптимизация моделей оценки вероятности дефолта в кризисных условиях

Влияние пандемии COVID-19 на пруденциальную основу МСФО (IFRS) 9

С точки зрения многих банков, снижение прогнозов экономического роста является предвестником роста вероятности дефолта по многим заемщикам, а также ожидаемых кредитных потерь (в т.ч. из-за падения стоимости залогов, которое в целом является результатом общего падения цен на активы).

И хотя большинство правительств по всему миру предоставляют пакеты прямой финансовой помощи домохозяйствам, малому бизнесу и более крупным корпорациям, многие банки также получают настоятельные рекомендации поддерживать свою кредитную деятельность на том же или даже более высоком уровне, чем до кризиса. По сути, во многих странах банки становятся агентами косвенной государственной помощи. Это оказывается возможным, поскольку одновременно многие банковские регуляторы пытаются смягчить негативное воздействие пандемии COVID-19 и связанной с ней глобальной экономической рецессии на финансовые учреждения путем предоставления доступа к дополнительным источникам ликвид-



Оптимизация моделей оценки вероятности дефолта в кризисных условиях

ности, а также временного послабления регуляторных нормативов и пруденциальной нагрузки¹.

Под действие подобных регуляторных послаблений попадает и подход к оценке ожидаемых кредитных потерь, устанавливаемый МСФО (IFRS) 9. Пионером в этой области выступил европейский регулятор ЕБА², выпустив «Рекомендации о применении пруденциальной основы в отношении определения дефолтов, реструктуризаций и принципов МСФО (IFRS) 9 в свете мер, принимаемых в связи с COVID-19»³. В данном документе ЕБА отмечает, что вводимые правительствами многих стран кредитные моратории не являются специфическими для отдельных заемщиков, а напротив, направлены на поддержку целых классов заемщиков в рамках конкретных банковских продуктов и отраслей. С учетом этого банкам предлагается не классифицировать соответствующие кредитные соглашения как реструктурированные и воздерживаться от автоматического включения их во вторую «корзину» обесценения на этом основании. Похожее по смыслу инструкции, разрешающие временно не ухудшать оценку кредитного качества заемщиков, имевших высокие показатели до начала пандемии, выпустил в марте–апреле и Банк России⁴.

С другой стороны, регуляторы подчеркивают, что подобные рекомендации не отменяют обязательств кредитных организаций адекватно оценивать свои риски и выявлять любые ситуации реального несоответствия заемщиков заданным критериям кредитного качества. В частности, ЕБА отмечает возросшую важность применения профессионального суждения и анализа качественных критериев при определении стадии обесценения активов⁵. В практическом аспекте это означает, что финансовые учреждения должны будут серьезно пересмотреть свои нынешние подходы к расчетам ожидаемых кредитных убытков по МСФО (IFRS) 9, сосредоточившись на тех областях, где быстрая адаптация позволит получить наилучший экономический эффект. Мы хотели бы выделить три таких области.

Под действие регуляторных послаблений попадает и подход к оценке ожидаемых кредитных потерь, устанавливаемый МСФО (IFRS) 9.

¹ Так, например, в рамках борьбы с кризисными последствиями пандемии Банк России, помимо прочего, принял мораторий на проведение некоторых видов проверок до 01.07.2020, предоставил послабления в части соблюдения системно значимыми банками норматива краткосрочной ликвидности Н26 (Н27), а также ввел новые инструменты по предоставлению банками дополнительной средне- и долгосрочной ликвидности в формате аукционов репо на срок 1 месяц и 1 год (cbr.ru/info_2020/).

² European Banking Authority.

³ Statement on the application of the prudential framework regarding Default, Forbearance and IFRS 9 in light of COVID 19 measures (eba.europa.eu).

⁴ См. Информационные письма Банка России от 17.04.2020 № ИН-01-41/72 и от 20.03.2020 № ИН-01-41/19.

⁵ "Institutions would be expected to distinguish between obligors for which the credit standing would not be significantly affected by the current situation in the long term, from those that would be unlikely to restore their credit worthiness. Exercising this discrimination would contribute to mitigate any potential cliff effect of transfers between stages and would help to avoid exaggerating the effects of the shock". Источник: eba.europa.eu.

Алексей АНТОНОВ
Роман СОРОКИН

Классификация кредитов по стадиям («корзинам») обесценения

Как было отмечено ранее, простые критерии значительного ухудшения кредитоспособности, которые широко применялись большинством банков в последние 2–3 года, такие как наличие просроченной задолженности со сроком более 30 дней, значительное ухудшение кредитного рейтинга, оценка вероятности дефолта или кредитного спреда по сравнению со значениями на дату первоначального признания актива, могут быть недостаточными.

В новых условиях при оценке кредитного качества заемщика значительно возрастает роль индивидуального анализа. Это означает, что набор критериев «значительного ухудшения кредитного качества» должен быть дополнен рядом количественных, а также (что более важно) качественных факторов, указывающих на способность бизнеса выживать и выполнять свои обязательства в среднесрочной и долгосрочной перспективе.

Набор критериев «значительного ухудшения кредитного качества» должен быть дополнен рядом количественных, а также (что более важно) качественных факторов, указывающих на способность бизнеса выполнять свои обязательства в среднесрочной и долгосрочной перспективе.

Динамическое прогнозирование ECL

Поскольку правительства и регулирующие органы будут стимулировать финансовые учреждения к продолжению кредитования, для банков все важнее будет возможность прогнозировать будущие значения и проводить стресс-тестирование ECL с использованием сложных сценариев. Эти сценарии должны учитывать актуальные бизнес-планы по кредитованию, макроэкономический стресс, увеличение доли активов второй стадии, рост числа реструктуризаций и дефолтов и т.д.

Оценка факторов риска (PD и LGD)

Большинство моделей PD и LGD построены на основе исторической информации, которая не может отразить специфику кризисного периода. Первым и очевидным шагом здесь являются актуализация модели макроэкономической поправки с использованием обновленных прогнозных данных в части PD-моделей, а также обновление параметров, связанных с оценкой стоимости обеспечения (таких как потеря стоимости обеспечения при реализации, расходы на реализацию объекта залога, а также время, необходимое для продажи).

Однако возможности по улучшению моделей этим не исчерпываются. Для риск-факторов, моделируемых на коллективной основе, интересные результаты может дать применение более точной и глубокой сегментации портфелей, а также альтернативных алго-

Оптимизация моделей оценки вероятности дефолта в кризисных условиях

ритмов калибровки наблюдаемых частот дефолтов к параметрическим кривым PD. Это может позволить получить более точные оценки PD с учетом характера экономической деятельности заемщика, размера фирмы, географии или любых других факторов.

Именно на последней области — пожалуй, наименее очевидной, но при этом достаточно перспективной — мы хотели бы сфокусироваться далее.

Оптимизация моделей вероятности дефолта, основанных на портфельном анализе

В отличие от индивидуального кредитного анализа, когда для моделирования вероятности дефолта используется целый спектр индивидуальных количественных и качественных параметров заемщика, характеризующих его кредитоспособность, портфельный анализ основан на предположении об однородности кредитного качества и уровня кредитного риска всех заемщиков, входящих в портфель. Соответственно основным фактором, используемым при оценке вероятности дефолта, является наблюдаемая частота дефолтов, или Observable Default Rate (ODR)¹. Рассчитывая ODR по соответствующей группе рейтинга, мы получаем типичное поведение для контракта в рамках портфеля. При этом общеизвестным фактом является необходимость сегментации кредитного портфеля банка для получения более однородных групп заемщиков.

Далее приведен пример ODR², рассчитанных в разрезе риск-классов (выделены на основании глубины просроченной задолженности и наличия ряда признаков, указывающих на «существенное ухудшение кредитного качества») и продуктовых сегментов. Очевидно, что разные продукты, а также принадлежность к определенному риск-классу напрямую влияют на вероятность дефолта. Данную закономерность можно увидеть на графиках ODR (рис. 1; period — время в месяцах от момента попадания в соответствующую группу).

Согласно нашим наблюдениям, во время процесса моделирования вероятности дефолта с использованием портфельного подхода многие банки ограничиваются статической сегментацией портфеля (обычно основанной на продукте, отраслевом секторе или типе контр-

При появлении новых наблюдений, характерных для кризисного периода, целесообразным становится применение моделей машинного обучения для более точной сегментации (которую может быть сложно получить на основании экспертных суждений).

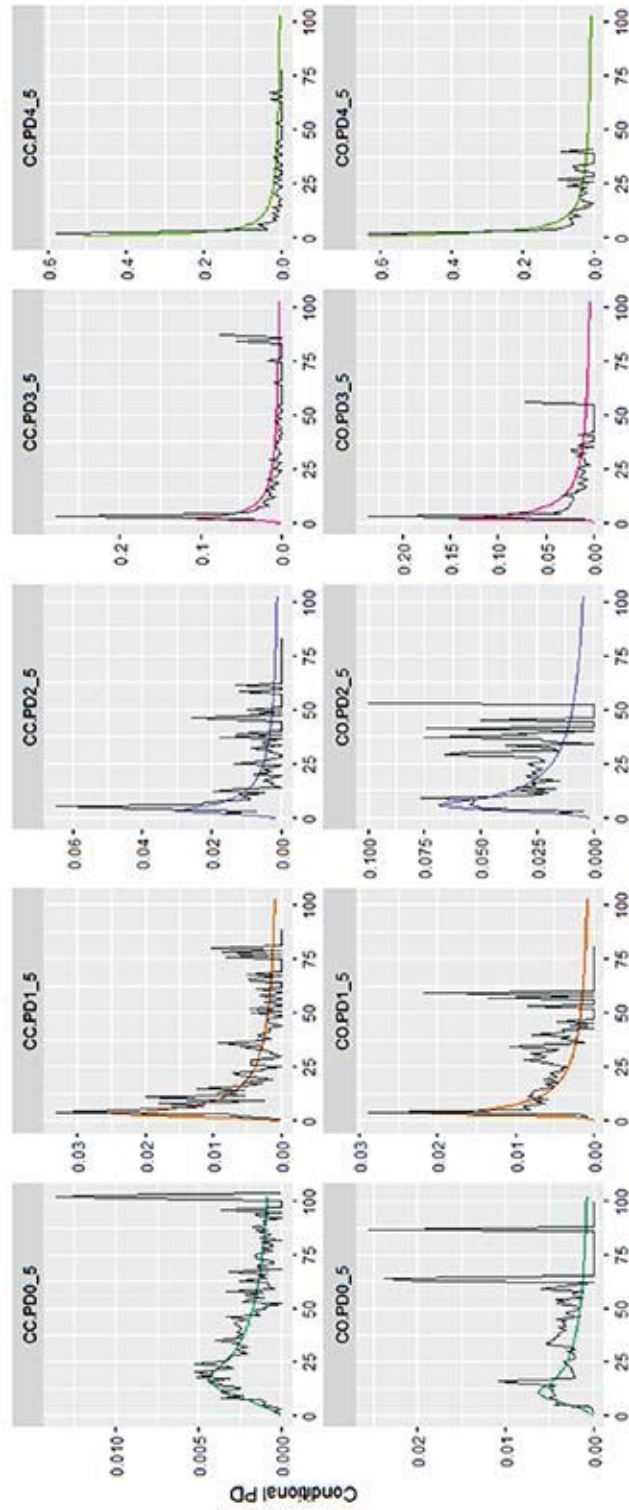
¹ Под ODR мы понимаем эмпирическую оценку условной вероятности дефолта или вероятность дефолта в момент времени t , если контракт выжил в момент времени $t-1$: $ODR_t = \frac{\text{Количество контрактов в дефолте}_t}{\text{Общее количество контрактов под риском}_t}$. В анализе выживаемости также встречается термин hazard.

² Источник: результаты моделирования вероятности дефолта для одного из банков — клиентов компании «Неофлекс».

Алексей АНТОНОВ
Роман СОРОКИН

Рисунок 1

ODR в разрезе риск-классов и кредитных продуктов (CO и CC)



Оптимизация моделей оценки вероятности дефолта в кризисных условиях

агента), полученной на основе экспертного суждения. Обладая хорошим балансом простоты и эффективности, данный подход, во многом оправдывавший себя в нормальных экономических условиях, может оказаться недостаточно эффективным в условиях кризиса. Действительно, ограничительные меры могут совершенно по-разному влиять, например, на различные компании (или их сотрудников) из одних и тех же отраслей, пользующихся одними и теми же кредитными продуктами. В качестве иллюстрации можно привести ИТ-отрасль: сотрудники компаний, специализирующихся на разработке средств удаленной работы, смогут увеличить свой доход и станут менее склонны к дефолтам или сохранят прежний уровень риска, а сотрудники компаний, специализирующихся, скажем, на обслуживании авиационной отрасли, начнут испытывать серьезные трудности.

Таким образом, при появлении новых наблюдений, характерных для кризисного периода, целесообразным становится применение моделей машинного обучения для более точной сегментации (которую может быть сложно получить на основании экспертных суждений). На практике достаточно хорошие результаты могут быть получены при использовании алгоритмов кластеризации или специальных деревьев решений для работы с категориальными признаками. В качестве примеров можно привести ряд библиотек, реализованных на R: partykit, CHAID, c50. Также перспективными, на наш взгляд, являются подходы, основанные на кластеризации временных рядов, где в качестве отдельного временного ряда рассматривается динамика остатка долга по кредиту.

Калибровка параметрических PD-кривых с использованием алгоритмов Auto ML

После завершения процедур кластеризации и расчета ODR целесообразно осуществить калибровку наблюдаемых частот дефолтов к параметрической кривой. Это позволит сгладить волатильность ODR и получить возможности расчета Lifetime PD для контрактов, срок действия которых превышает статистику наблюдений. На практике в нормальных экономических условиях достаточно хорошим выбором является распределение Вейбулла-Парето без дополнительных признаков. Данная функция распределения хорошо описывает наличие большего количества дефолтов на ранних сроках жизни контрактов портфеля и уменьшающееся количество дефолтов по мере приближения к дате погашения, что согласуется с теоретическими предположениями:

Алексей АНТОНОВ
Роман СОРОКИН

$$F(t) = A \times \left(1 - e^{-\alpha \left(\frac{t}{t_0} \right)^k} \right) \times \theta(t_0 - t) + \left(1 - B \times \left(\frac{t_0}{t} \right)^\beta \right) \times \theta(t - t_0),$$

где $\theta(t)$ — функция Хевисайда;

$k > 0$;

$A, B, \beta > 1$;

t_0 — точка перехода распределения Вейбулла в распределение Парето.

Непрерывность $F(t)$ и ее первой и второй производных в t_0 обеспечивает фиксацию параметров α , A и B соотношением $u = \beta/k$.

Параметры α , β , k оцениваются с помощью метода максимального правдоподобия (LM):

$$L(\theta | \text{data}) = \prod_{i \in D} f(T_i; \theta) \prod_{i \in R} S(C_i; \theta) \rightarrow \max_{\theta}.$$

При этом:

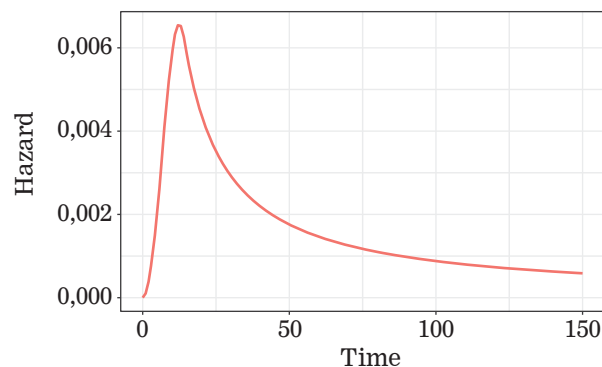
- если i находится в дефолте, наблюдение i вносит вклад $f(T_i)$;
- если i цензурирована, наблюдение i вносит вклад $Pr(T > C_i) = S(C_i)$.

На рис. 2 приведен результат аппроксимации ODR для сегмента потребительских кредитов, группы риск-класса PD0_9 (кредиты без просроченной задолженности) распределением Вейбулла–Парето.

Однако в кризисных условиях, подобных текущим, графики ODR могут принимать нетипичный вид — например, получить несколько пиков либо пики или высокую волатильность на длительных сроках контрактов. Это может наложить серьезные ограничения на качество

Рисунок 2

Пример: результат аппроксимации ODR для сегмента потребительских кредитов



Оптимизация моделей оценки вероятности дефолта в кризисных условиях

аппроксимации PD-кривой при использовании распределения Вейбулла-Парето. Для решения данной проблемы может быть использован подход в парадигме Auto ML, который обеспечивает автоматический выбор модели, наилучшим способом описывающей имеющиеся данные о наблюдаемых частотах дефолтов.

В нашем случае для решения этой задачи на языке R была специально реализована надстройка для разного типа моделей: интересные нас PD-модели с помощью «обертки» приводятся к общему стандарту, использующему методы `fit` и `predict`. Независимо от особенностей реализации внутреннего кода и алгоритма моделей входные и выходные данные будут одинаковыми. «Обертки» для моделей реализованы в классах R6, что позволяет создать общий класс доступных моделей `RiskModels`, удобный для дальнейшего сравнения и выбора. Для включения новой модели в `RiskModels` достаточно написать скрипт-обертку по простому примеру.

Поскольку основной задачей Auto ML является валидация доступных моделей, также был реализован оптимизатор, включающий следующие методы:

- отбор признаков;
- сэмплирование данных для кросс-валидации;
- бэк-тест;
- сравнение моделей.

Итогом работы оптимизатора является набор обученных моделей, сэмплов кросс-валидации, прогнозов, рассчитанных метрик и графиков.

Говоря об отборе признаков, мы подразумеваем следующие возможные подходы:

1. Перебор всех комбинаций указанных признаков для модели.
2. Предусмотренная L1-регуляризация для логистической регрессии.
3. Понижение размерности (PCA-анализ, PLS-регрессия).
4. Отбор признаков с исключением (`recursive feature elimination` или `permutation feature importance`).
5. Также для указанных признаков применяется биннинг (на основе WOE или специальных деревьев решений).

В рамках такого подхода обеспечивается возможность применения по сути неограниченного количества альтернативных алгоритмов моделирования PD-кривой без необходимости «ручной» оценки и сравнения полученных результатов. При этом мы можем использовать не только функции распределения без дополнительных признаков (как упомянутое выше распределение Вейбулла-Парето), но

Модель пропорционального риска Кокса, логистическая регрессия и обобщенное гамма-распределение оказались «финалистами», понятными аудиторам банка и устойчиво проходящими валидацию.

Алексей АНТОНОВ Роман СОРОКИН

и другие алгоритмы, способные обучаться на разных наборах признаков: как категориальных, так и количественных. В этих случаях статичными признаками могут выступать, например, различные характеристики клиента, такие как скорбалл на момент выдачи кредита, сфера занятости и т.п. В качестве динамических признаков могут быть использованы информация о просроченной задолженности за последние N месяцев, доля выплаченной суммы кредита на дату наблюдения, история транзакций клиента, макроэкономические индикаторы и многое другое. Далее будет рассмотрен пример работы подобных моделей, включающих категориальные и макроэкономические данные.

В нашем примере, выполненном на данных одного из банков — клиентов нашей компании, мы использовали модель пропорционального риска Кокса, логистическую регрессию и обобщенное гамма-распределение. Однако набор используемых моделей не обязательно должен быть ограничен указанными выше: в нашем случае они оказались «финалистами», понятными аудиторам банка и устойчиво проходящими валидацию. Так, изначально в нашу модель был включен и ряд других алгоритмов: Age-Period-Cohort-анализ, Random Forest, Random Forest Survival и т.д. В нашем случае их применение не показало хороших результатов (в частности, Random Forest склонен к переобучению), поэтому мы исключим их из дальнейшего обзора. Отметим, что на выбранном портфеле не лучшим образом показала себя и модель Кокса, однако это скорее исключение из правил. Ввиду высокой популярности этой модели, ее фундаментального значения в анализе выживаемости и нашего опыта применения этой модели было решено также проиллюстрировать работу модели в рамках данной статьи.

В моделях, описанных ниже, в качестве количественного признака применялась наиболее часто используемая макроэкономическая переменная — индекс динамики ВВП. В качестве примера использования категориальных признаков использовался срок выдачи кредита¹.

Далее мы рассмотрим некоторые модели более детально для дальнейшего сравнения². Для большей наглядности вместо вывода коэффициентов для каждой модели будут продемонстрированы графики зависимостей признаков с итоговой оценкой условной вероятности дефолта.

¹ Для выбранного портфеля на уровне генерации признаков статистически было выявлено, что кредиты разделены по сроку выдачи до двух лет и более.

² Параметры распределения Вейбулла-Парето были описаны выше.

Оптимизация моделей оценки вероятности дефолта в кризисных условиях

Обобщенное гамма-распределение с признаком контрактного срока¹

Удобно использовать представление через функцию распределения:

$$f(t) = \frac{|\beta|}{\Gamma(k) \times t_0} \times \left(\frac{t}{t_0}\right)^{k\beta-1} \times e^{-\left(\frac{t}{t_0}\right)^\beta},$$

где $t_0 > 0$ — параметр масштаба;

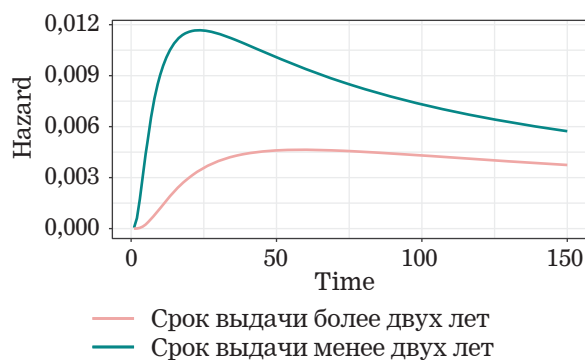
$\beta > 0, k > 0$ — параметры формы;

$\Gamma(k)$ — гамма-функция, которая определяется как $\Gamma(x) = \int_0^\infty s^{x-1} \times e^{-s} ds$.

Hazard оценивается с помощью обобщенного гамма-распределения с учетом срока выдачи контракта. Эффект от разделения контрактов по сроку выдачи наглядно представлен на рис. 3.

Рисунок 3

Результаты модели обобщенного гамма-распределения с признаком контрактного срока



Модель пропорционального риска Кокса с признаками контрактного срока и динамики ВВП

Существует большое количество модификаций данного подхода, чтобы сделать его «непропорциональным»². Мы рассмотрим только базовую версию алгоритма³ и его применения. Предполагается, что функция риска представима в следующем виде:

¹ Stacy E.W. A Generalization of the Gamma Distribution // The Annals of Mathematical Statistics. 1962. Vol. 33. No. 3. P. 1187-1192.

² Гушин К., Козлов А. Новый подход к оценке поведенческой вероятности дефолта // Риск-менеджмент в кредитной организации. 2017. № 3. С. 35-42.

³ Cox D.R. Regression Models and Life-Tables // Journal of the Royal Statistical Society. Series B (Methodological). 1972. Vol. 34. No. 2. P. 187-220.

Алексей АНТОНОВ
Роман СОРОКИН

$$h(t, x_1, x_2(t)) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2(t)),$$

где $h_0(t)$ — базовая функция риска (baseline hazard), когда все остальные ковариаты равны нулю. В данном случае мы будем оценивать ее с помощью распределения Вейбулла–Парето (для каждого набора данных необходимо отдельно оценивать «базовый риск» — это могут быть различные распределения или сглаживающие методы);

β_i — оцениваемые неизвестные параметры модели;

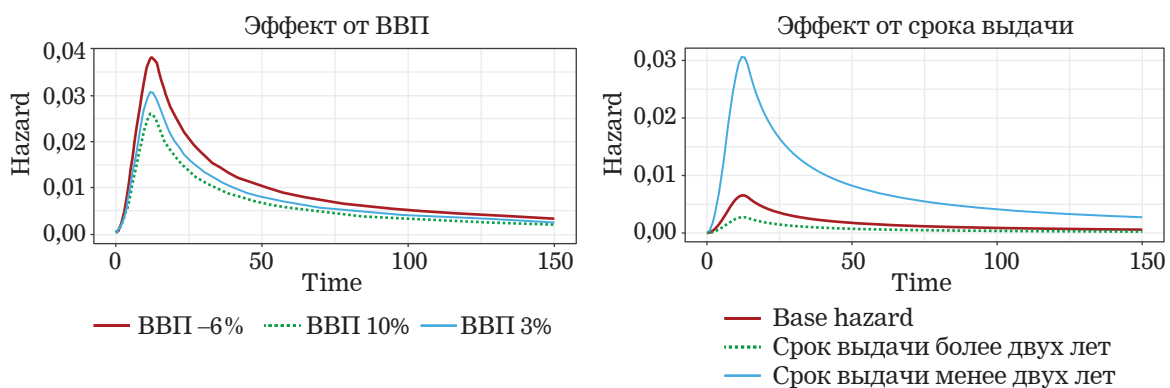
x_1 — переменные, характеризующие клиента, например сфера деятельности заемщика, скорбалл на момент выдачи кредита и т.д. В рассматриваемой модели x_1 — это бинарная переменная по сроку выдачи кредита менее или более двух лет;

x_2 — макроэкономический индекс роста ВВП (в условиях COVID-19 также стоит рассмотреть возможность включения как минимум индекса безработицы).

Эффект от включения в модель Кокса соответствующих объясняющих переменных наглядно представлен на рис. 4.

Рисунок 4

Результаты модели пропорционального риска Кокса с признаками контрактного срока и динамики ВВП



Модель логистической регрессии с признаками контрактного срока и динамики ВВП

Модель может быть записана в виде:

$$h(X) = \frac{1}{1 + \exp -(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4)},$$

где β_i — оцениваемые неизвестные параметры модели;

x_1 — бинарная переменная по сроку выдачи кредита менее или более двух лет;

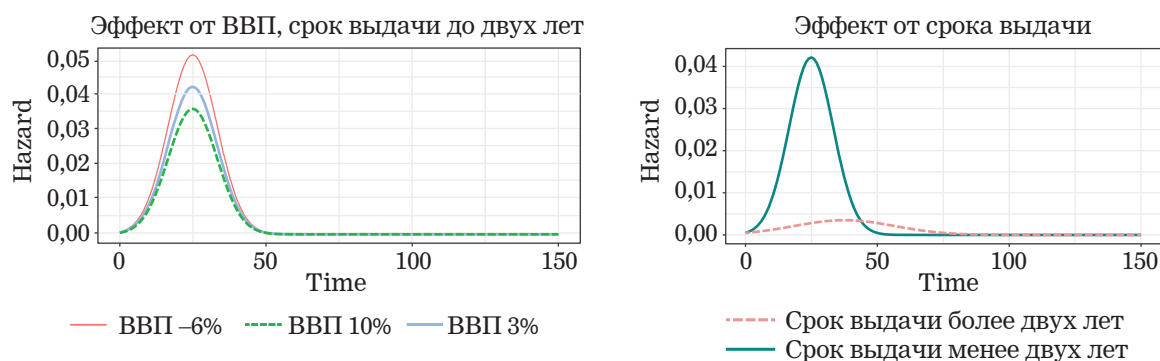
Оптимизация моделей оценки вероятности дефолта в кризисных условиях

x_2 — макроэкономический индекс динамики ВВП;
 x_3 — месяц жизни контракта (Month of Book, MOB);
 x_4 — квадрат месяца жизни контракта (MOB²).

Соответствующие модельные PD-кривые представлены на рис. 5.

Рисунок 5

Результаты модели логистической регрессии с признаками контрактного срока и динамики ВВП



В отличие от пропорциональной модели Кокса логистическую регрессию можно легко сделать непропорциональной, то есть будет допускаться разная временная зависимость вероятности дефолта для контрактов с разными сроками выдачи кредитов. Однако к недостаткам логистической модели можно отнести быстро убывающий до нуля правый хвост — это хорошо видно на приведенных графиках.

Бэк-тестинг полученных кривых

После инициализации моделей мы выполняем автоматизированную проверку на пригодность к дальнейшему использованию на данных тестовой выборки.

В качестве тестовой выборки можно взять данные за последний год как наиболее актуальные (рис. 6).

Предполагается, что обученные модели должны пройти бэк-тест (рис. 7) на тестовой выборке. Для проверки адекватности прогноза мы использовали односторонний биномиальный критерий. На основе статистического критерия строятся доверительные интервалы и оценивается частота пробития¹. Детальное описание процедуры выходит за рамки данной статьи.

¹ Битюцкий В., Патратий О., Перевицкая В. Методика моделирования достаточности капитала: валидация / Ассоциация российских банков, Комитет по стандартам Базеля II и управлению рисками, 2013 (arb.ru/b2b/docs/validatsiya_komitet_arb_po_standartam_bazel_ii_i_upravleniyu_riskami-9752430/).

Алексей АНТОНОВ
Роман СОРОКИН

Рисунок 6

Деление данных для обучения и тестирования

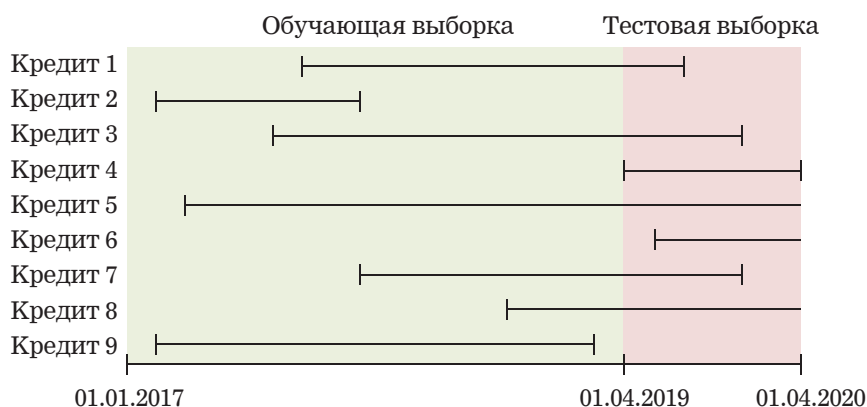
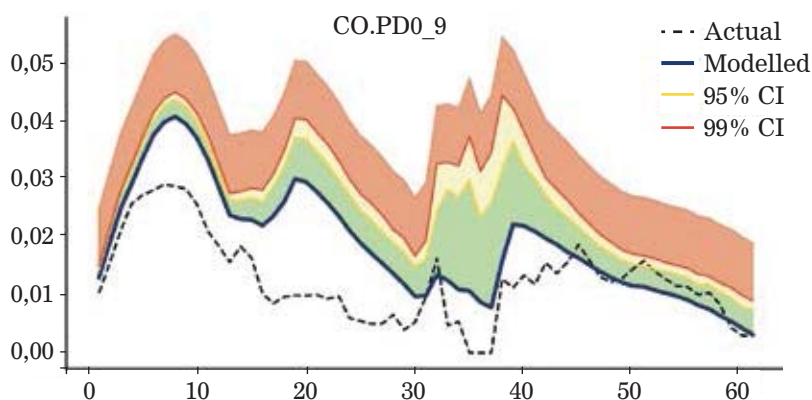


Рисунок 7

Визуализация процедуры бэк-теста



Модели, прошедшие бэк-тест, принимаются для дальнейшей оптимизации. На следующем шаге прогнозы моделей на тестовой выборке сравниваются по заранее выбранной метрике и в результате выбирается модель, наиболее точно описывающая реальные данные. Стоит обратить внимание, что типичные для скоринга метрики, такие как ROC AUC, Gini и т.д., затруднительны для интерпретации в случае Lifetime PD, поскольку при моделировании вероятности дефолта на протяжении всей жизни кредита мы получаем целую кривую, а не вероятность дефолта как метку для контракта.

Оптимизация моделей оценки вероятности дефолта в кризисных условиях

Выбор оптимальной PD-кривой

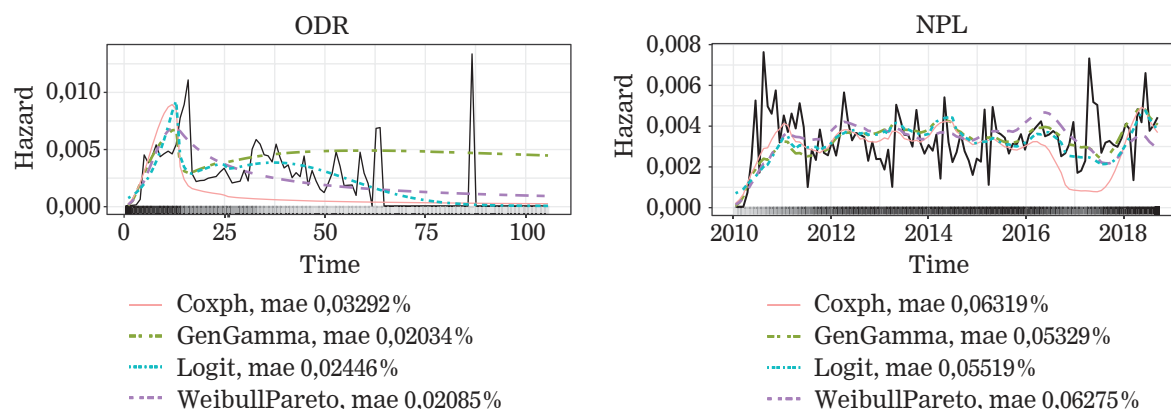
Технически для оценки качества моделей можно использовать регрессионные метрики, сравнивая прогнозные и фактические значения. В зависимости от целей, преследуемых кредитной организацией, можно применить один из трех подходов:

1. Оценка фактических и модельных ODR.
2. Оценка фактических и модельных потерь по кредитам (объем NPL¹).
3. Выбор оптимального размера модельных резервов.

В первом и втором случаях мы получаем результаты, которые представлены на рис. 8.

Рисунок 8

Сравнение фактических и модельных NPL или ODR



Для сравнения NPL с модельными прогнозами мы будем использовать стандартную метрику Mean Absolute Error (MAE). Для сравнения по ODR предлагается использовать MAE, взвешенную по доле контрактов в данный месяц жизни кредитов ($wMAE$), поскольку с увеличением наблюдаемого срока жизни кредита количество наблюдаемых кредитов уменьшается и значения ODR могут принимать аномальные значения. Формула расчета:

$$wMAE = \frac{\sum_{i=1}^N |\text{ODR_real}_i - \text{ODR_fit}_i| \times W_i}{n},$$

где t — номер месяца жизни кредита;

W_i — доля контрактов для i -го месяца жизни кредита;

ODR_fit — модельная оценка, взвешенная по всем контрактам.

¹ Non-Performing Loans — наблюдаемая доля дефолтов относительно календарного времени.

Алексей АНТОНОВ Роман СОРОКИН

Использование данной метрики в значительной степени будет оправдано в том случае, если банк будет стремиться получить набор PD-кривых, максимально точно описывающих фактическое поведение наблюдаемых дефолтов.

Аналогично можно сравнивать модельные оценки резервов и фактические убытки. Рассмотрим одну из возможных метрик для этого подхода, предусматривающую возможность штрафов с разными весами в зависимости от знака ошибки:

$$\text{Loss}(t) = \sum_{i:r_i > q_t}^{N1} \alpha \times (r_i - q_t)^2 + \sum_{j:r_j < q_t}^{N2} \beta \times (r_j - q_t)^2,$$

где $t \in \{t_1, \dots, t_i, \dots, t_{N1}, t_{N1+1}, \dots, t_j, \dots, t_{N2}\}$ — номера месяца;
 r_t — фактический размер кредитных потерь;
 q_t — прогнозный размер кредитных потерь;
 α — штрафной коэффициент за недостаточность резервов;
 β — штрафной коэффициент за избыточные резервы.

Коэффициенты α и β могут быть откалиброваны с учетом аппетита банка и его текущих приоритетов (консервативный подход к формированию резервов или, наоборот, стремление к их снижению). Фактические кредитные потери r_t могут быть получены из исторических данных, а модельные резервы q_t — оценены по упрощенной формуле:

$$q_t = \text{EAD}_t \times \text{MPD}_t \times \text{LGD}_t,$$

где EAD_t — сумма остатков всех кредитов на балансе банка плюс объем внебалансовых контрактов, поправленный на фактор кредитной конверсии (CCF);

MPD_t — вероятность дефолта по сегменту в месяце t ;

LGD_t — средний уровень возможных потерь в случае дефолта по сегменту.

После того как по всем группам риска выбранного сегмента рассчитаны значения метрик, можно перейти к сравнительному анализу и выбору лучшей модели (рис. 9).

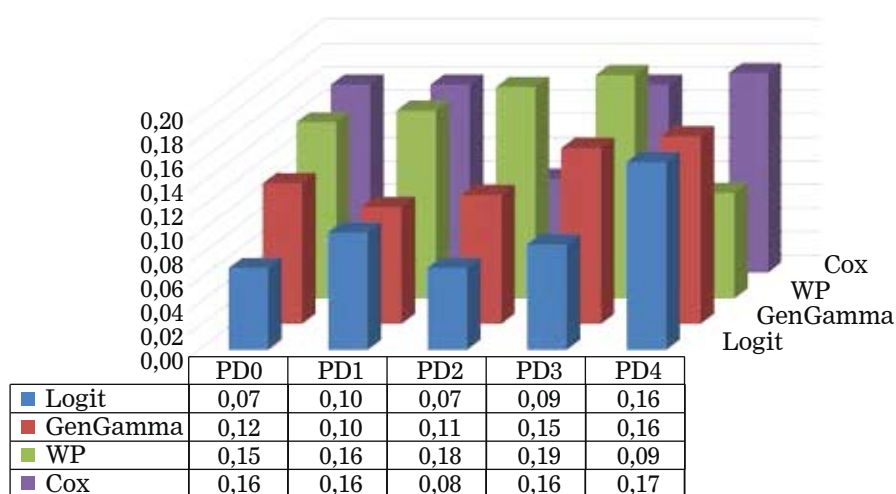
Мы видим, что наименьшие ошибки в прогнозировании наблюдаются при использовании модели логистической регрессии. Однако в риск-классе PD0_4 (в нашем примере это контракты с DPD¹ 60+ дней) лучшей моделью является распределение Вейбулла–Парето. Выбирать ли лучшую модель в среднем для всех групп или исполь-

¹ Days Past Due — количество дней просрочки на указанную дату.

Оптимизация моделей оценки вероятности дефолта в кризисных условиях

Рисунок 9

Сравнение ошибок моделей на основе метрики оценки NPL



зывать для каждой группы просрочки свою лучшую модель — этот вопрос оставляется на усмотрение риск-аналитика.

Наконец, если банк заинтересован в минимизации резервов (что в текущих условиях представляется менее актуальным), целесообразно применять подход, основанный на выборе кривых, обеспечивающих возможность получения минимальных значений ECL. Однако для этого требуется более точное моделирование резервов, которое может быть выполнено с использованием либо итерационных расчетов в автоматизированной системе расчета резервов, либо приближенной оценочной формулы, учитывающей разбивку контрактов сегмента n по стадиям обесценения:

$$ECL^i = \sum_{t=1}^{12} EAD_t^{\text{Stage 1}} \times MPD_t^i \times LGD_t + \sum_{t=1}^T EAD_t^{\text{Stage 2}} \times MPD_t^i \times LGD_t,$$

где i — сценарий кредитного риска (совокупность PD-кривых, рассчитанных по конкретному алгоритму);

Stage — стадия обесценения (первая или вторая, т.к. контракты в дефолте нет смысла включать в данное сравнение);

T — срок жизни самого длинного контракта в месяцах.

По нашему опыту, применение подобного подхода к оптимизации резервов может приносить вполне ощутимые практические результаты. Так, для одного из банков, где мы применили данный подход, сокращение модельных резервов составило примерно 7% по отно-

Алексей АНТОНОВ
Роман СОРОКИН

шению к объему резервов, сформированных с использованием стандартной модели.

Уже сейчас многие финансовые институты столкнулись с необходимостью значительно увеличить объем резервов из-за пересмотра критериев присвоения стадий обесценения и актуализации макроэкономических поправок с учетом наиболее свежих прогнозов.

Например, уже в первом квартале 2020 г. банк HSBC увеличил прогнозный объем ECL на 14%, Barclays — на 21%, а Сбербанк — на 10%¹. Поэтому сегодня актуальность задачи построения более точных моделей PD и LGD для целей МСФО (IFRS) 9 сложно переоценить. Концепция Auto ML позволяет автоматизировать сегментацию портфеля, feature engineering и выбор оптимального алгоритма для калибровки PD-кривой с учетом бэк-тестинга (что крайне важно с точки зрения прохождения банками внешнего аудита), текущих задач и риск-аппетита банка. Также данный подход позволит значительно сократить трудозатраты риск-менеджеров на сопровождение подобных моделей и заложит основу для их постоянной адаптации к меняющимся экономическим условиям. 

¹ По данным квартальной отчетности финансовых организаций (находится в открытом доступе).