ОЦЕНКА СОСТОЯНИЯ РОЗНИЧНОГО КРЕДИТНОГО ПОРТФЕЛЯ НА ОСНОВЕ КОНСТРУКЦИЙ ИЗ ПАРНЫХ КОПУЛ

Ратников А.А., Щетинин Е.Ю.

Финансовый университет при правительстве Российской Федерации,
Россия, г. Москва, Ленинградский пр-т, д.53/1
aa_ratnikov@mail.ru, riviera-molto@mail.ru

В настоящее время банки активно развиваются в сфере розничного кредитования. Характерной особенностью розничных продуктов является необходимость использования продвинутого математического аппарата на всех стадиях процесса: от генерации лидов и выдачи кредита, до последующего мониторинга выданных кредитов. Процесс мониторинг пула выданных банком кредитов принято называть управлением кредитного портфеля. Одной из ключевых задач в управлении кредитным портфелем является оценка риска дефолта заёмщиков, то есть оценка вероятности того, что клиент на одном из платежей выйдет на просрочку. От адекватности и актуальности оценки риска напрямую зависит эффективность кредитной политики банка.

Ключевые слова: вьющиеся копулы, оценка риска,

Введение

Сложность задачи прогнозирования кредитного портфеля заключается в том, что на вероятность дефолта влияют множество зависящих друг от друга параметров заёмщика и предлагаемого кредитного продукта. Пренебрежение взаимосвязями параметров может привести к сильному искажению оценок риска дефолта, а учёт взаимосвязей порождает сложную многомерную структуру данных.

В данной работе рассматривается подход оценки кредитного портфеля на основе аппарата вьющихся копул (vine copulas). Такой подход позволяет свести задачу к совокупности парных копул, для которых задача поиска закона распределения и оценки параметров становится значительно проще. Основные вычисления проводятся на языке R с использованием пакетов CDVine и VineCopula.

В последние несколько лет в научной литературе наблюдается увлечение конструкциями из копул при исследовании эффектов, в основе которых лежат случайные величины. Данную тенденцию можно объяснить общим увеличением доступных для анализа данных и усложнением основанных на них математических моделей. Становится очевидным, что в прикладных задачах данные часто имеют сложные внутренние нелинейные зависимости, пренебрежение которыми может привести к серьезным ошибкам в оценках и расчетах.

Математические конструкции, основанные на копулах позволяют частично решить проблемы, связанными с зависимостью переменных. В частности, метод вьющихся копул позволяет разбить сложную многомерную структуру на иерархический набор парных копул, удобных для последующих анализа и оценки.

В рамках данной работы будет рассмотрен общий подход к построению и оценки парных и вьющихся D-копул с помощью пакета CDVine и VineCopula. В качестве примера использования конструкций из парных копул будет продемонстрирован подход к задаче оценки риска дефолтности портфеля розничных кредитов.

1 Конструкции из парных копул

Пусть d-мерный случайный вектор с совместной функцией распределения и с одномерными маргинальными функциями распределения x_d). Копулой будем называть функцию, которая позволяет представить совместное распределение в терминах мартингалов и структуры их зависимости. Теорема Шкляра утверждает, что копула, связанная с F является d-мерной функцией распределения $C: [0,1]^d \rightarrow [0,1]$, удовлетворяющая равенству:

(1)

где $\boldsymbol{\beta}$ является векторным параметром зависимостей между мартингалами. Для удобства записи в дальнейшем обычно будем скрывать данный параметр.

Определим плотность копулы, как:

(2) $$c(F_1, \dots, F_d) = \frac{\delta C(F_1, \dots, F_d)}{\delta F1, \dots, F_d}$$

Для абсолютно непрерывной, строго возрастающей F можно получить уравнения совместной функции плотности вероятности f вида:

(3) $$f(x_1, \dots, x_d) = c_{1\dots d}\big(F_1(x_1), \dots, F_d(x_d)\big) \cdot \prod_{k=1}^{d} f_k(x_k)\,,$$

где $c_{1\dots d}(\cdot)$ однозначно определённая плотность d-мерной копулы $C_{1\dots d}$.

Рассмотрим в качестве примера гауссову копулу. В общем виде для d-мерного случайного вектора $\boldsymbol{X}$ она может быть записана как:

(4) $$C_X = \Phi_X\big(\Phi^{-1}(x_1), \dots, \Phi^{-1}(x_d)\big),$$

где Φ^{-1} – обратная функция распределения нормального закона, а Φ_X совместная функция распределения нормального закона с нулевым математическим ожиданием и корреляционной матрицей R

Матрица R в данном случае выступает в качестве многомерного параметра, отражающего связь мартингалов. Обычно для упрощения процедуры оценки параметров корреляционную матрицу R сводят к единственному параметру ρ такому что:

(5) $$R = \begin{pmatrix} 1 & \dots & \rho \\ \vdots & \ddots & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

Параметр ρ в гауссовой копуле отражает степень зависимости между мартингалами.

Отметим, что на практике использование копул высокой размерности для оценки и моделирования некоторого случайного вектора нецелесообразно из-за проблемы «проклятия размерности» и общей сложности и многопараметричности функции $\mathrm{C}_{1\dots d}(\cdot)$.

При соблюдении условий регулярности d-мерная функция распределения f может быть итерационно выражена через парные копулы, на основе общей формулы условной функции распределения:

(6) $$f(x_k|\boldsymbol{\nu}) = c_{x_k,\nu_j|\boldsymbol{\nu}_{-j}}\Big(F\big(x_k\big|\boldsymbol{\nu}_{-j}\big), F\big(\nu_k\big|\boldsymbol{\nu}_{-j}\big)\Big) \cdot f(x_k|\boldsymbol{\nu}_{-j}),$$

где $\nu = \boldsymbol{x}_k$ есть вектор размерности d-1, не включающий в себя некоторую компоненту x_k, а $\boldsymbol{\nu}_{-j}$ – соответственно вектор ν, не включающий в себя компоненту ν_j.

Очевидно, что итерационный процесс, отраженный в данной формуле, может привести к различным вариантам разложения f на множества условных парных копул.

Разложение многомерной копулы на множество парных копул называется ветвлением. Для визуализации ветвления используется граф, вершины которого соответсвуют одномерным случайным величинам, а рёбра -- параметрам связи между ними. Таким образом любые две вершины, связанные ребром, интерпретируются как одна парная копула. Обычно в литературе выделяют три типа ветвления: R, C, D. В рамках данной работы подробно рассмотрим D-ветвление.

Общую формулу плотности вероятности для D-ветвление можно записать как:

(7) $$f(x_k; \phi) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{i=1}^{d-1}\prod_{j=1}^{d-i} c_{j,j+i|(j+1):(j+i-1)}\Big(F\big(x_j\big|x_{j+1}, \dots, x_{j+i-1}\big), F\big(x_{j+i}\big|x_{j+1}, \dots, x_{j+i-1}\big)\Big)$$

В качестве примера можно получить разложение для трёхмерного случайного вектора $\boldsymbol{X}$:

(8) $$f(x_{18}, x_2, x_3; \boldsymbol{\phi}) = \prod_{k=1}^{3} f_k(x_k) \times c_{12}\big(F_1(x_1), F_2(x_2)\big)\, c_{23}\big(F_2(x_2), F_3(x_3)\big) c_{13|2}\big(F(x_1|x_2), F(x_3|x_2)\big),$$

где $\boldsymbol{\phi} = (\boldsymbol{\beta}_{12}, \boldsymbol{\beta}_{23}, \boldsymbol{\beta}_{13|2})$.

2 Моделирование и оценка D-вьющихся копул

Для демонстрации использования аппарата вьющихся копул проведём численный эксперимент. С помощью функции CDVineSim() пакета CDVine, реализованном на языке R, выполним симуляцию трёхмерной D-вьющийся копулы с параметрами: C_{12} – Гауссова копула, C_{23} – Гауссова копула, $C_{13|2}$ – копула Клайтона, $\theta_{12} = 0.2$, $\theta_{23} = 0.3, \theta_{13|2} = 1$.

Корреляционная матрица полученных данных имеет вид:

$$R(x,y,z) = \begin{pmatrix} 1 & 0.197 & 0.814 \\ 0.197 & 1 & -0.02 \\ 0.814 & -0.02 & 1 \end{pmatrix} \tag{9}$$

Решим задачу восстановления распределения на основе данных, полученных в результате симуляции. Данную задачу будем решать в два шага: определим законы распределения на основе информационного критерия Акаике (AIC); уточним значение параметров копул с помощью метода максимального правдоподобия.

Указанные шаги выполняются при помощи функций CDVineCopSelect(selectioncrit = "AIC") и CDVineMLE().

В результате первого шага были корректно определены виды копул (две Гауссовы копулы и копула Клайтона), а также получены предварительные оценки параметров: (0.248, 0.31, 1.03).

В ходе второго шага оценки параметров были скорректированы: (0.24, 0.28, 1.04).

Отметим, что полученные оценки близки к изначально заданным параметрам.

Наделим данные прикладным смыслом. Для этого переопределим одну из трех переменных как функцию-индикатор по формуле:

$$Z' = \begin{cases} 0, & if\ z < 0.5 \\ 1, & if\ z \geq 0.5 \end{cases} \tag{10}$$

Новая выборка будет иметь смысл некоторого кредитного розничного портфеля, где каждая точка соответствует одному выданному кредиту. Случайные величины X, Y определяют некоторые нормированные характеристики клиента или кредита (займа), например сумму кредита (займа), скоринговый бал, долговую нагрузку и т.д. Величина Z' говорит о факте выхода клиента в дефолт.

Для того, чтобы определить уровень дефолтности в пространстве признаков клиента/кредита (займа), то есть вероятности того, что клиент с таким набором признаком уйдет в просрочку, для каждой точки рассчитывается среднее значение величин z' для k ближайших соседей в пространстве признаков. Расстояние будем определять, как евклидово.

Полученную выборку можно оценить как (n+1) мерную D-ветвящуюся копулу, где n - число исследуемых параметров.

На основе определенных видов парных копул и их оцененных параметров генерируем новую выборку с помощью функции CDVineSim().

Предложенный подход позволяет оценивать распределение вероятности дефолта на любых срезах признаков, обходя проблему «проклятия размерности». Полученные в результате симуляции данные могут быть использованы для выявления возникающих позитивных и негативных трендов в портфеле, а также для ретро-тестирования внедряемых изменений в риск-политике банка.

Заключение

В работе был подробно рассмотрен аппарат конструкций из парных копул, в частности D-вьющиеся копулы, для задачи оценки многомерных случайных величин. В практической части работы была продемонстрирована генерация, оценка и симуляция на основе полученной оценки трехмерной D-вьющейся копулы с помощью пакета CDVine в R. Также был предложен и продемонстрирован подход для оценки уровня дефолтности кредитного портфеля. Предложенный подход имеет потенциал в задах оценки кредитного риска, ретро-тестирования внедряемых изменений в риск-политике банка, а также выявлении позитивных и негативных трендов в изменении уровня риска и входящего клиентского потока.

Литература

1. *Sklar, A.*, 1959. Fonctions de ŕepartition `a n dimensions et leurs marges.Publications de l'Institut de Statistique de l'Universit́e de Paris 8, 229–231

2. *Kim D. et al.* Mixture of D-vine copulas for modeling dependence //ComputationalStatistics Data Analysis. – 2013. – Т. 64. – С. 1-19.

3. *Shchetinin E.Yu.*, В сборнике: Extreme copulas modeling with maxstablespatial processes, Распределенные компьютерные и телекоммуникационные се-ти: управление, вычисление, связь (DCCN-2017) Материалы Двадцатой международной научной конференции. под общ. ред. В.М. Вишневского. 2017. С.495-502.

4. *Shchetinin E.Yu., Martynova V.M.,* Precipitation maximum spatial dependencestructures modeling, Всборнике: CEUR Workshop Proceedings 7. Сер. "SelectedPapers from the Proceedings of the 7th International Conference "Informationand Telecommunication Technologies and Mathematical Modeling of High-TechSystems ITTMM 2017"2017. С. 72-77.
5. *Коновалов К.А., Щетинин Е.Ю.,* Иерархические типы структур статистиче-ских зависимостей, Вестник Российского университета дружбы народов. Серия: Математика, информатика, физика, 2009, 3, С. 68-71.
6. *Акимов В.А., Быков А.А., Щетинин Е.Ю.,* ВВЕДЕНИЕ В СТАТИСТИКУ ЭКСТРЕМАЛЬНЫХ ЗНАЧЕНИЙ И ЕЕ ПРИЛОЖЕНИЯ. МЧС России; Всероссийский научно-исследовательский институт по проблемам гражданскойобороны и чрезвычайных ситуаций МЧС России. Москва, 2009.