ТЕОРІЯ ОПТИМАЛЬНИХ РІШЕНЬ

УДК 519.85

В.В. БОЙКО, В.Н. КУЗЬМЕНКО

ОБ ЭФФЕКТИВНОСТИ ВЫЧИСЛЕНИЯ НЕКОТОРЫХ ВИДОВ ОШИБОК В ТЕОРИИ РИСК-КВАДРАТОВ

Рассматривается альтернативный вариант определения некоторых видов ошибок и потерь в теории риск-квадратов. Показано, что для рассматриваемого варианта ошибки и потери можно вычислять более эффективно, используя лишь один оптимизируемый параметр.

Введение. В работе [1] построена теория риск-квадратов (Risk Quadrangle) – наборов функций, определенных на случайной величине. В каждый набор входит четыре связанные между собой функции – риск, отклонение, потери и ошибка (risk, deviation, regret, error). Функции связаны между собой общими для любого набора соотношениями. При выполнении определенных условий риск-квадраты называются регулярными и позволяют определять те или иные характеристики случайной величины путем минимизации функций ошибки или потерь, а в некоторых случаях путем минимизации риска и отклонения.

Приведем основные определения и связи между функциями, входящими в регулярный риск-квадрат.

Пусть X – случайная величина, $E(X)$ – математическое ожидание X, $V(X)$ – некоторая функция, называемая «потери». Тогда функция ошибки по определению равна $\mathcal{E}(X)=V(X)-E(V)$, риск определяется как $R(X)=\min_B\{B+V(X-B)\}$, отклонение как $D(X)=\min_B \mathcal{E}(X-B)$, статистика – $S(X)=$
$= \arg\min_B \mathcal{E}(X-B) = \arg\min_B\{B+V(X-B)\}$.

Справедливы следующие равенства: $V(X)=\mathcal{E}(X)+E(X)$, $R(X)=D(X)+E(X)$.

Регулярный риск квадрат может быть определенным образом отмасштабирован.

Линейная комбинация регулярных риск-квадратов также может быть регулярным риск-квадратом.

Выше, как и в [1], функция риска определяется через функцию потерь. Приведем теорему о смешивании риск-квадратов и покажем, что для некоторых риск-квадратов можно ввести обратное определение, а именно, определить потери через риск, одновременно упростив задачу оптимизации, которую нужно решать для вычислений. (Для упрощения записи символ случайной величины X будем по возможности опускать.)

Теорема о смешивании риск-квадратов [1]. Пусть $(R_k, D_k, V_k, \varepsilon_k)$, $k = 1,\dots,K,$ есть регулярные риск-квадраты со статистиками S_k, а для весов $\lambda_k > 0$ выполнено $\lambda_1 + \lambda_2 + \dots + \lambda_K = 1.$ Тогда некий риск-квадрат (R, D, V, ε) со статистикой S задается соотношениями:

$$S(X) = \lambda_1 S_1(X) + \dots + \lambda_K S_K(X),$$
$$R(X) = \lambda_1 R_1(X) + \dots + \lambda_K R_K(X),$$
$$D(X) = \lambda_1 D_1(X) + \dots + \lambda_K D_K(X),$$
$$V(X) = \min_{B_1,\dots,B_K} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k V_k(X - B_k) \,|\, \sum\nolimits_{k=1}^{K} \lambda_k B_k = 0 \right\},$$
$$\varepsilon(X) = \min_{B_1,\dots,B_K} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k \varepsilon_k(X - B_k) \,|\, \sum\nolimits_{k=1}^{K} \lambda_k B_k = 0 \right\}.$$

Наша цель – это упрощение процедур вычисления функций потерь $V(X)$ и ошибки $\varepsilon(X)$. Обе процедуры – эквивалентны, в том смысле, что $\varepsilon(X) = V(X) - EX.$

Рассмотрим вычисление $V(X)$ как прямую задачу оптимизации

$$V(X) = \min_{B_1,\dots,B_K} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k V_k(X - B_k) \,|\, \sum\nolimits_{k=1}^{K} \lambda_k B_k = 0 \right\}. \quad (1)$$

Значение $V(X)$ может быть найдено как решение такой минимаксной задачи:

$$V(X) = \min_{B_1,\dots,B_K} \max_{\eta} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k V_k(X - B_k) + \eta \sum\nolimits_{k=1}^{K} \lambda_k B_k \right\} =$$
$$= \max_{\eta} \min_{B_1,\dots,B_K} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k \left(\eta B_k + V_k(X - B_k) \right) \right\} =$$
$$= \max_{\eta} \left\{ \sum\nolimits_{k=1}^{K} \lambda_k \min_{B_k} \left(\eta B_k + V_k(X - B_k) \right) \right\}. \quad (2)$$

Заметим, что минимизация в (2) «похожа» на минимизацию в определении риска $R(X),$ но включает дополнительный параметр $\eta.$ $R(X)$ определяется так:

$$R(X) = \min_{B} (B + V(X - B)). \quad (3)$$

В работе [1] приведены примеры различных риск-квадратов. Часть риск-квадратов содержит некоторый параметр $a,$ а функции потерь могут быть представлены в виде $V(X,a)=\phi(a)\psi(X)$, где $\phi(a)$ – некоторая функция от параметра a.

В этом случае минимизация в (2) может быть представлена в виде

$$\min_B(\eta B+V(X-B))=\min_B(\eta B+\varphi(a)\psi(X-B))=\eta\min_B(B+\frac{\varphi(a)}{\eta}\psi(X-B)). \quad (4)$$

Если для каждого значения $\eta\neq 0$ существует такое значение параметра a_η, что $\varphi(a_\eta)=\varphi(a)/\eta$, то последнее выражение в (4) можно записать как:

$$\eta\min_B(B+\varphi(a_\eta)\psi(X-B))=\eta\min_B(B+V(a_\eta,X-B))=\eta R(X,a_\eta).$$

Для каждого конкретного квадрата область изменения η имеет свои связанные со свойствами $R(X,a_\eta)$ ограничения, а случай $\eta=0$ анализируется отдельно.

Если $K=1$, то определение (2) становится таким:

$$V(X,a)=\max_\eta\left(\eta R(X,a_\eta)\right). \quad (5)$$

Заметим, что в противоположность к (3), где риск находится путем минимизации выражения, включающего потери, в (5) потери находятся путем максимизации выражения, включающего риск.

В случае, когда в смесь входят одинаковые функции с разными значениями параметров, (1) может быть представлено в виде

$$V(X,\vec{a})=\max_\eta\left\{\eta\sum_{k=1}^K\lambda_k R_k(X,a_{\eta k})\right\}, \quad (6)$$

где $\vec{a}=(a_1,...,a_K)$ – вектор параметров.

Тогда, когда функция потерь (ошибки) может быть напрямую вычислена, формулы (5) и (6) имеют лишь теоретический интерес как констатация некоторого свойства риск-квадратов с параметрами. Но когда в (1) действительно нужна оптимизация и для вычисления риска $R_k(X,a_{\eta k})$ есть «быстрый» алгоритм, то использование (6) дает существенно более простую одномерную задачу оптимизации и делает процесс вычисления (например, минимизации ошибки) более эффективным.

Приведем примеры, в которых фунции потерь могут быть определены через риск. Все примеры, кроме последнего, взяты из [1]. Последний – из [2].

Предварительно дадим необходимые определения и обозначения.

Для параметра $\alpha\in(0,1)$ будем рассматривать такие характеристики X: α-квантиль $q_\alpha(X)$ и α-суперквантиль $Q_\alpha(X)$.

α-квантиль $q_\alpha(X)=\sup\{x\,|\,F_X(x)<\alpha\}$, может обозначаться как $VaR_\alpha(X)$ (value-at-risk), где $F_X(x)=P(X\le x)$ – функция распределения X.

α-суперквантиль $Q_\alpha(X)=\int_\alpha^1 q_\beta(X)d\beta$ определяется как среднее значение X для значений не меньших, чем $q_\alpha(X)$ и может обозначаются как $CVaR_\alpha(X)$ (conditional value-at-risk).

Номера примеров и названия риск-квадратов сохранены те же, что и в [1].

Пример 2 [1]. Квантильный (Quantile) риск-квадрат. Статистика $S(X,\alpha) = = q_\alpha(X)$, риск $R(X,\alpha) = Q_\alpha(X)$, потери $V(X,\alpha) = \dfrac{1}{1-\alpha} E[\max\{0, X\}]$.

Альтернативное определение $V(X,\alpha) = \max\limits_\eta \eta Q_{\alpha_\eta}(X)$, $\alpha_\eta = 1-\eta(1-\alpha)$.

Пример 5 [1]. Пессимистический (Worst Case) риск-квадрат. Статистика $S(X) = \sup(X)$, риск $R(X) = \sup(X)$, потери $V(X) = \begin{cases} 0 \ if \ X \le 0 \\ \infty \ if \ X > 0 \end{cases}$.

Альтернативное определение $V(X) = \sup\limits_\eta \eta \sup(X)$.

Пример 6 [1]. Распределенный пессимистический (Distributed Worst Case) риск-квадрат.

Статистика $S(X) = p_1 \sup(X_1) + ... + p_r \sup(X_r)$;

риск $R(X) = p_1 \sup(X_1) + ... + p_r \sup(X_r)$;

потери $V(X) = \begin{cases} 0 \ if \ p_1 \sup(X_1) + ... + p_r \sup(X_r) \le 0 \\ \infty \ if \ p_1 \sup(X_1) + ... + p_r \sup(X_r) > 0 \end{cases}$.

Альтернатива $V(X) = \sup\limits_\eta \eta\left(p_1 \sup(X_1) + ... + p_r \sup(X_r)\right)$.

Здесь $p_1,...,p_r$ вероятности некоторых условий такие, что $p_1 + ... + p_r = 1$, $X_1,...,X_r$ – условные распределения случайной величины X. Должно быть выполнено условие $p_1 EX_1 + ... + p_r EX_r = EX$.

Пример 10 [1]. Смешанный квантильный (Mixed Quantile) риск-квадрат.

Статистика $S(X,\vec{\alpha}) = \lambda_1 q_{\alpha_1}(X) + ... + \lambda_r q_{\alpha_r}(X)$, $(\lambda_i > 0,\ \lambda_1 + ... + \lambda_r = 1)$;

риск $R(X,\vec{\alpha}) = \lambda_1 Q_{\alpha_1}(X) + ... + \lambda_r Q_{\alpha_r}(X)$;

потери $V(X,\vec{\alpha}) = \min\limits_{B_1,...,B_r} \left\{ \sum_{k=1}^r \lambda_k V_k(X - B_k, \alpha_k) \mid \sum_{k=1}^r \lambda_k B_k = 0 \right\}$,

$V_k(X - B_k, \alpha_k) = \dfrac{1}{1-\alpha_k} E[\max\{0, X - B_k\}]$.

Альтернатива $V(X,\vec{\alpha}) = \max\limits_\eta \eta \sum_{k=1}^r \lambda_k Q_{\alpha_{\eta k}}(X)$, $\alpha_{\eta k} = 1-\eta(1-\alpha_k)$.

Пример 12 [1]. Высшего порядка ($p > 1$) квантильный (Higher Order Quantile) риск-квадрат.

Статистика $S(X,\alpha) = q_\alpha^{(p)}(X)$;

риск $R(X,\alpha) = Q_\alpha^{(p)}(X)$;

потери $V(X,\alpha) = \dfrac{1}{1-\alpha}\left(E[\max\{0, X\}]^p\right)^{\frac{1}{p}}$.

Альтернативное определение $V(X,\alpha)=\max_{\eta}\eta Q_{\alpha_\eta}^{(p)}(X)$, $\alpha_\eta=1-\eta(1-\alpha)$.

Пример из [2]. Суперквантильный (Superquantile) риск-квадрат.

Статистика $S(X,\alpha)=Q_\alpha(X)$;

риск $R(X,\alpha)=\dfrac{1}{1-\alpha}\int_\alpha^1 Q_\beta(X)d\beta$;

потери $V(X,\alpha)=\dfrac{1}{1-\alpha}\int_0^1 \max\{0,Q_\beta(X)\}d\beta$.

Альтернативное определение $V(X,\alpha)=\max_{\eta}\eta\dfrac{1}{1-\alpha_\eta}\int_{\alpha_\eta}^1 Q_\beta(X)d\beta$,

$\alpha_\eta=1-\eta(1-\alpha)$.

Говоря об альтернативном способе определения и вычисления функции потерь, мы имеем ввиду и вычисление функции ошибки, поскольку ошибка легко вычисляется через потери $\mathcal{E}(X)=V(X)-EX$.

Формально параметр η не ограничен при оптимизации в (2). Но в примерах, содержащих суперквантиль $Q_{\alpha_\eta}(X)$, значение η должно быть ограничено интервалом $0<\eta<1/(1-\alpha)$, для которого $\alpha_\eta=1-\eta(1-\alpha)\in(0,1)$. Покажем, что оп-тимальное значение η в (2) лежит в замыкании этого интервала для кван-тиль-ного риск-квадрата. Для этого задачу минимизации в (2) представим в виде $v(\eta,X)=\inf_B\left(\eta B+\dfrac{1}{1-\alpha}E[\max\{0,X-B\}]\right)$ и найдем зависимость решения от η.

Для $\eta<0$ $v(\eta,X)=-\infty$ при $B\to+\infty$.

Для $\eta=0$ $v(\eta,X)=0$ при $B\to+\infty$.

Для $\eta\to 0^+$ $v(\eta,X)=0$ при $B\to+\infty$.

Для $0<\eta<\dfrac{1}{1-\alpha}$ $v(\eta,X)=\eta Q_{\alpha_\eta}(X)$ при $B=q_{\alpha_\eta}(X)$.

Для $\eta=\dfrac{1}{1-\alpha}$ $v(\eta,X)=\dfrac{1}{1-\alpha}EX$ при $B=-\infty$.

Для $\eta>\dfrac{1}{1-\alpha}$ $v(\eta,X)=-\infty$ при $B=-\infty$.

Отсюда следует, что для поиска максимума в (5), достаточно рассмотреть интервал $0 < \eta < \frac{1}{1-\alpha}$ и его граничные значения. Аналогично для остальных примеров, содержащих $Q_{\alpha_\eta}(X)$. При поиске максимума для смеси квантилей по формуле (6) интервал равен $0 < \eta < \frac{1}{1 - \min\limits_k \alpha_k}$.

Выводы. Показано, что для регулярных риск-квадратов существуют различные способы определения функций потерь и ошибок. Для определения, предложенного в данной работе, функции потерь и ошибок для некоторых риск-квадратов могут быть посчитаны через риск. Такое определение может дать существенное упрощение вычисления функций ошибок, например, в случае использования смеси квантилей [3].

В.В. Бойко, В.М. Кузьменко

ПРО ЕФЕКТИВНІСТЬ РОЗРАХУНКУ ДЕЯКИХ ВИДІВ ПОХИБОК
У ТЕОРІЇ РИЗИК-КВАДРАТІВ

Розглядається альтернативний варіант визначення деяких видів похибок та втрат у теорії ризик-квадратів. Показано, що для варіанта, який пропонується, похибки та втрати можливо розраховувати більш ефективно, використовуючи лише один параметр, що оптимізується.

V.V. Boyko, V.M. Kuzmenko

ABOUT EFFECTIVNESS OF CALCULATION OF SOME ERRORS IN RISK-QUDRANGLE THEORY

An alternative variant of defining some errors and regrets in the risk-quadrangle theory is considered. It is shown than errors and regrets may be calculated more effectively using only one optimized parameter for proposed variant.

1. *Rockafellar R.T., Uryasev S.* The Fundamental Risk Quadrangle in Risk Management, Optimization, and Statistical Estimation. *Surveys in Operations Research and Management Science*. **18**. 2013. P. 33 – 53.
2. *Rockafellar R.T., Royset J.O., Miranda S. I.* Superquantile regression with applications to buffered reliability, uncertainty quantification and conditional value-at-risk. *European J. Operations Research*. **234**. 2014. P. 140 – 154.
3. *Кузьменко В.Н., Ненахов Э.И.* О взаимозаменяемости некоторых функций ошибки при регрессионном анализе. *Теорія оптимальних рішень*. К.: Ін-т кібернетики імені В.М. Глушкова НАН України, 2016. С. 129 – 136.

Получено 01.03.2017