

УДК 330.46, 519.254, 519.812.3

СРАВНЕНИЕ СИСТЕМ УПРАВЛЕНИЯ ИНВЕСТИЦИОННЫМИ ПОРТФЕЛЯМИ ПРИ МНОГОКРИТЕРИАЛЬНОЙ ОЦЕНКЕ КАЧЕСТВА

В.Г. Саркисов, Г.А. Саркисов

Самарский государственный технический университет
Россия, 443100, г. Самара, ул. Молодогвардейская, 244

Авторы продолжают исследование стабильности качества работы систем управления инвестиционными портфелями на финансовом рынке, начатое в [1]. Рассмотрены системы управления портфелем, оптимизируемые в рамках многокритериальной задачи. Предложен метод перекрестного ретроспективного тестирования, позволяющий оценить снижение качества работы системы, вызванное неоптимальностью настроек, а также сравнить эффективность различных систем при параметрической оптимизации и тестировании на множестве выборок. Представлен ряд вариантов доминирования отображений эффективных границ множеств допустимых портфелей для сравнения различных оптимизированных систем. Описана иерархия вариантов такого доминирования.

Ключевые слова: перекрестная проверка, многокритериальная оптимизация, параметрическая оптимизация, инвестиционный портфель, эффективная граница.

Проблема потери эффективности системами управления инвестиционными портфелями стала особенно актуальной на фоне относительно быстрых изменений конъюнктуры мировых финансовых и товарных рынков. В подобных условиях особенно важным является создание систем управления, способных сохранять качество работы при изменении свойств временных рядов цен финансовых инструментов.

Для проверки качества работы и оптимизации систем [2] используется ретроспективное тестирование – тестирование системы на прошлом материале. Обычно исторический временной ряд цен делится на две части (выборки): обучающую и тестовую.

На обучающей выборке система строится и оптимизируется. Упрощенно можно сказать, что в ходе оптимизации из всего множества рассматриваемых систем со всевозможными наборами параметров выбирается подмножество наиболее эффективных по заданным критериям (рис. 1).

На тестовой выборке проверяется качество работы отобранных систем. При сравнении систем обычно выбирается та, которая показала лучшую эффективность на тестовой выборке.

Операции, проводимые на обеих выборках, различаются лишь множествами рассматриваемых систем и методами оптимизации, но не принципами отбора. Более того, можно сказать, что на тестовой выборке осуществляется оптимизация путем перебора систем.

Заметим, что решение о том, какая именно из двух выборок будет обучаю-

Виген Геннадьевич Саркисов (к.т.н.), доцент кафедры «Высшая математика и прикладная информатика».

Геннадий Арсенович Саркисов (к.т.н., доц.), доцент кафедры «Высшая математика и прикладная информатика».

щей, а какая – тестовой, в общем случае основывается на случайных факторах (например, хронологической последовательности) и не направлено непосредственно на выбор более качественной системы. Более того, если поменять местами обучающую и тестовую выборки, то в большинстве случаев изменится и выбор лучшей системы.



Рис. 1. Схема отбора систем и параметров при традиционном подходе

Многокритериальность постановки задачи оптимизации усугубляет проблему неоднозначности выбора.

Также отметим, что системы, показавшие наилучшие результаты на обучающей выборке, в большинстве случаев показывают плохие результаты на тестовых. Это явление называется перепоподгонкой (overfitting) и связано с излишне точной аппроксимацией обучающей выборки с помощью системы. Обычно эти системы характеризуются сложной разветвленной структурой алгоритма и большим числом настраиваемых параметров.

Интерес представляет подход к тестированию, который позволил бы выявить системы, способные получать стабильно хорошие результаты на различных временных отрезках. Подобный подход для случая однокритериальной постановки задачи представлен авторами в [1]. В настоящей работе полученные ранее результаты распространяются на случай многокритериальной оценки качества работы системы и соответствующей задачи ее многокритериальной оптимизации.

Перекрестная проверка системы при однокритериальной параметрической оптимизации

Рассматриваются системы управления инвестиционными портфелями, содержащие вектор настраиваемых параметров. Вектор параметров i -й системы обозначим через $a_i = (a_{i,1}, a_{i,2}, \dots, a_{i,N_i})$, где N_i – количество параметров i -й системы.

Основой подхода к сравнению систем является перекрестная проверка – параметрическая оптимизация («обучение») исследуемых систем на различных сочетаниях нескольких из K имеющихся выборок и анализ качества их работы с оптимизированными параметрами на остальных (тестовых) выборках. При этом количество используемых для каждого обучения выборок может быть выбрано в пределах от 1 до $K-1$. Слишком большое количество выборок в обучающем наборе приводит к тому, что оптимальные параметры, получаемые при оптимизации на разных наборах, практически идентичны получаемым на объединении всех имеющихся выборок. Слишком малое количество выборок в обучающем

наборе, наоборот, приводит к существенному разнообразию векторов параметров. Исходя из этих соображений при исследовании стабильности целесообразно производить обучение на малых наборах обучающих выборок, а при формировании системы для практического использования – на больших наборах.

Далее при исследовании стабильности используется обучение по одной выборке. Критерий качества управления портфелем с помощью i -й системы на k -й выборке обозначим $C_1(a_i, k)$.

Пусть, например, при наличии K выборок i -я система оптимизируется («обучается») на первой из них:

$$a_i^{*1} = \arg \max_{a_i} C_1(a_i, 1). \quad (1)$$

С полученным вектором параметров a_i^{*1} i -я система тестируется на оставшихся выборках (вычисляются $C_1(a_i^{*1}, 2), \dots, C_1(a_i^{*1}, K)$). Затем производится оптимизация на второй выборке (находится a_i^{*2}) и тестирование на остальных (вычисляются $C_1(a_i^{*2}, 1), C_1(a_i^{*2}, 3), \dots, C_1(a_i^{*2}, K)$) и так далее. В результате будет получена квадратная матрица оценок качества системы, причем номер строки матрицы будет соответствовать номеру обучающей выборки, а номер столбца – номеру тестовой. При однокритериальной оптимизации элементами матрицы являются значения критерия качества управления портфелем i -й системой:

$$R_i = \left\{ \overbrace{\begin{pmatrix} C_1(a_i^{*1}, 1) & \dots & C_1(a_i^{*1}, K) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*K}, 1) & \dots & C_1(a_i^{*K}, K) \end{pmatrix}}^{\text{тестовые выборки}} \right\} \text{обучающие выборки}. \quad (2)$$

Сопоставление полученных матриц R_i для разных систем может служить основой для сравнения качества самих систем. Если для всех $k, l = 1 \dots K$ $C_1(a_i^{*k}, l) > C_1(a_j^{*k}, l)$, то очевидно, что i -я система предпочтительнее j -й. Если же такое преимущество наблюдается лишь для части элементов матриц, то необходимо ввести дополнительное правило для сравнения систем. Для построения этого правила можно, например, рассмотреть критерий качества как случайную величину. Сравнение функций распределения или числовых характеристик таких случайных величин позволит выбрать более предпочтительную систему.

Перекрестная проверка при многокритериальной оптимизации

Многокритериальный подход к оценке качества управления портфелем усложняет реализацию приведенного выше подхода: в подавляющем большинстве случаев не существует доминирующего вектора параметров, который бы по всем критериям ($C_1 \dots C_M$) обеспечивал превосходство над системой с любым другим вектором параметров. Пример критериального множества с доминирующим вектором параметров a_i^{*1} (в двухкритериальной задаче) представлен на рис. 2.

Вектор параметров будем называть недоминируемым, если не существует ни одного вектора (доминирующего), который бы обеспечивал рассматриваемой системе качество управления портфелем по всем критериям не хуже данного и хотя бы по одному критерию лучше. На обучающей выборке может быть построено множество недоминируемых векторов (эффективная граница множества векторов). На рис. 3 представлено критериальное множество с несколькими недоминируемыми векторами ($a_i^{*1,1}$, $a_i^{*1,2}$, $a_i^{*1,3}$ и $a_i^{*1,4}$).

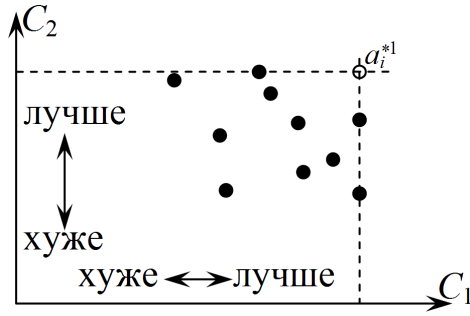


Рис. 2. Доминирующий вектор параметров

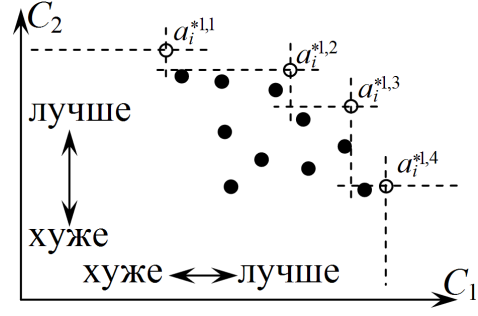


Рис. 3. Эффективная граница множества векторов параметров

Таким образом, существенным отличием перекрестной проверки при многокритериальной оптимизации является то, что каждая обучающая выборка порождает не один оптимальный вектор параметров, а множество недоминируемых векторов. Соответственно, проверять на тестовых выборках нужно будет каждый вектор этого множества.

Результаты тестирования требуют иного подхода к их интерпретации, так как для каждой пары обучающей и тестовой выборки необходимо проанализировать множество векторов параметров, каждый из которых характеризуется множеством критериев. В этом случае аналогом матрицы (2) является матрица (3) той же размерности $K \times K$. Однако элементами матрицы (3) будут уже не отдельные значения критерия, а матрицы значений критериев, в которых номер столбца соответствует номеру критерия $m = 1 \dots M$, а номер строки – номеру недоминируемого (на k -й обучающей выборке) вектора параметров $s = 1 \dots S_{ki}$:

$$R_i = \begin{pmatrix} \begin{pmatrix} C_1(a_i^{*1,1}, 1) & \dots & C_M(a_i^{*1,1}, 1) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*1, S_{1i}}, 1) & \dots & C_M(a_i^{*1, S_{1i}}, 1) \end{pmatrix} & \dots & \begin{pmatrix} C_1(a_i^{*1,1}, K) & \dots & C_M(a_i^{*1,1}, K) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*1, S_{1i}}, K) & \dots & C_M(a_i^{*1, S_{1i}}, K) \end{pmatrix} \\ \vdots & \ddots & \vdots \\ \begin{pmatrix} C_1(a_i^{*K,1}, 1) & \dots & C_M(a_i^{*K,1}, 1) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*K, S_{Ki}}, 1) & \dots & C_M(a_i^{*K, S_{Ki}}, 1) \end{pmatrix} & \dots & \begin{pmatrix} C_1(a_i^{*K,1}, K) & \dots & C_M(a_i^{*K,1}, K) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*K, S_{Ki}}, K) & \dots & C_M(a_i^{*K, S_{Ki}}, K) \end{pmatrix} \end{pmatrix}, \quad (3)$$

где

$$\begin{array}{cc} \text{обучающая} & \text{недоминируемый} \\ \text{выборка} & \text{вектор} \\ \hline C_m(a_i^{*k,s}, l) \\ \hline \text{критерий} & \text{тестовая} \\ & \text{выборка} \\ & \text{система} \end{array}$$

Внутренние матрицы, находящиеся в разных строках основной матрицы, могут иметь разное количество строк, так как количество недоминируемых векторов на разных обучающих выборках может различаться. Количество столбцов у всех внутренних матриц одинаковое и равно M .

Необходимо сформировать правило, позволяющее сравнить эту матрицу с аналогичной матрицей для другой системы и сделать соответствующие выводы о стабильности качества их работы.

Особенности оценок стабильности при многокритериальной оптимизации

Стабильность работы системы определяется сравнением элементов (т. е. внутренних матриц) столбца большой матрицы с элементом на пересечении этого столбца и главной диагонали. Это сравнение показывает, насколько система с оптимальными параметрами с других выборок уступает системе с оптимальными параметрами на данной выборке. Таким образом, при всех $k = 1 \dots K$, $k \neq l$ нужно

сравнить множество матриц $\begin{pmatrix} C_1(a_i^{*k,1}, l) & \dots & C_M(a_i^{*k,1}, l) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*k,S_{ki}}, l) & \dots & C_M(a_i^{*k,S_{ki}}, l) \end{pmatrix}$ с векторами пара-

метров, оптимизированных на k -х выборках, с матрицей $\begin{pmatrix} C_1(a_i^{*l,1}, l) & \dots & C_M(a_i^{*l,1}, l) \\ \vdots & \ddots & \vdots \\ C_1(a_i^{*l,S_{li}}, l) & \dots & C_M(a_i^{*l,S_{li}}, l) \end{pmatrix}$, у которой векторы параметров оптимизированы

на l -й выборке (l -я выборка одновременно является еще и тестовой).

У этих матриц одинаковое количество столбцов (по числу критериев), но разное количество строк (по числу недоминируемых векторов параметров на k -й и l -й выборках).

Пример графического представления рассматриваемых матриц для случая двух критериев ($M=2$) и трех выборок ($K=3$) при тестировании на первой выборке ($l=1$) приведен на рис. 4. Очевидно, что при тестировании на l -й выборке векторы $a_i^{*k,s}$, полученные при оптимизации на k -й выборке ($k \neq l$), могут быть недоминируемыми только в случае их совпадения с некоторыми из недоминируемых векторов $a_i^{*l,s}$ l -й выборки. В случае доминирования векторов $a_i^{*k,s}$ интерес представляет количественное описание превосходства над ними набора векторов $a_i^{*l,s}$.

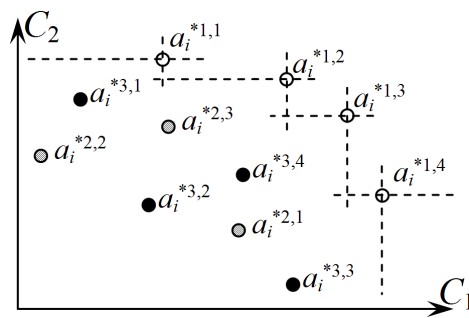


Рис. 4. Примеры результатов оптимизации и тестирования (в двухкритериальной задаче)

Рассматриваемая формулировка приводит к необходимости введения количественной меры различия не только для пары векторов критериев, но и для пары множеств таких векторов.

Такой подход имеет смысл в том случае, когда нужна оценка именно ста-

бильности безотносительно других свойств системы. Для сравнения систем между собой целесообразно сразу сравнивать порождаемые ими матрицы вида (3).

Сравнение матриц значений критериев для разных систем

Рассмотрим частные случаи, позволяющие без введения дополнительных связей и соотношений между критериями сравнить матрицы R_i и R_j вида (3), порождаемые i -й и j -й системами, и сделать вывод о предпочтительности i -й системы по сравнению с j -й.

1. Все элементы матрицы R_i являются доминирующими по Парето (по всем критериям (4) не хуже, хотя бы по одному (5) критерию m_0 лучше) по отношению ко всем элементам R_j , полученным при тестировании на тех же выборках l , кроме случая совпадения тестовой и обучающей выборок (рис. 5):

$$C_m(a_i^{*k_1, s_1}, l) \geq C_m(a_j^{*k_2, s_2}, l), \forall m, l, k_1 \neq l, k_2 \neq l, s_1, s_2; \quad (4)$$

$$\forall l, k_1 \neq l, k_2 \neq l, \exists m_0, s_1, s_2 : C_{m_0}(a_i^{*k_1, s_1}, l) > C_{m_0}(a_j^{*k_2, s_2}, l). \quad (5)$$

2. Все элементы матрицы R_i являются доминирующими по Парето по отношению ко всем элементам R_j , полученным при оптимизации и тестировании на тех же выборках, кроме случая совпадения тестовой и обучающей выборок (рис. 6):

$$C_m(a_i^{*k, s_1}, l) \geq C_m(a_j^{*k, s_2}, l), \forall m, l, k \neq l, s_1, s_2; \quad (6)$$

$$\forall l, k \neq l, \exists m_0, s_1, s_2 : C_{m_0}(a_i^{*k, s_1}, l) > C_{m_0}(a_j^{*k, s_2}, l). \quad (7)$$

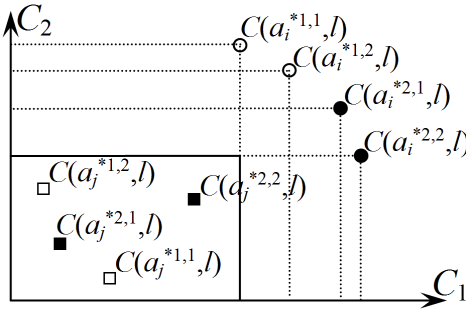


Рис. 5. Доминирование (вариант 1)

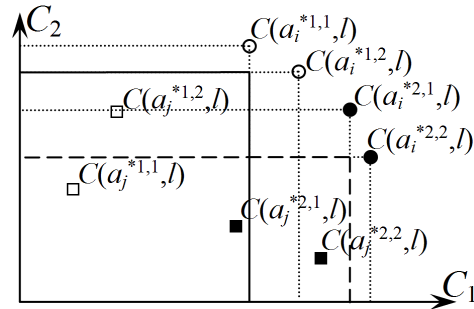


Рис. 6. Доминирование (вариант 2)

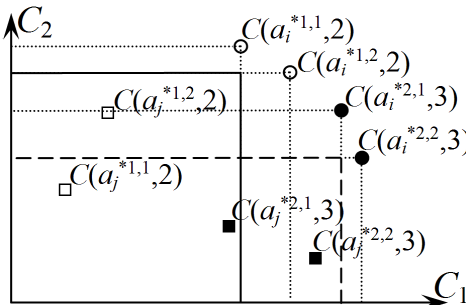


Рис. 7. Доминирование (вариант 3)

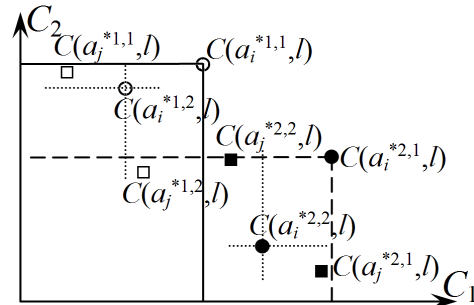


Рис. 8. Доминирование (вариант 4)

3. Третий вариант отличается от предыдущего тем, что рассматриваются пары обучающих и тестовых выборок, следующих в хронологической последовательности (рис. 7):

$$C_m(a_i^{*k,s_1}, k+1) \geq C_m(a_j^{*k,s_2}, k+1), \forall m, k < K, s_1, s_2; \quad (8)$$

$$\forall k < K, \exists m_0, s_1, s_2 : C_{m_0}(a_i^{*k_1,s_1}, k+1) > C_{m_0}(a_j^{*k_2,s_2}, k+1). \quad (9)$$

4. Элементы R_i являются недоминируемыми по Парето (10), а хотя бы один – доминирующим по Парето (10) по отношению ко всем элементам R_j , полученным на той же тестовой выборке, кроме случая совпадения тестовой и обучающей выборок (рис. 8):

$$\forall l, k_1 \neq l, k_2 \neq l, s_1, s_2 \exists m_0 : C_{m_0}(a_i^{*k_1,s_1}, l) > C_{m_0}(a_j^{*k_2,s_2}, l); \quad (10)$$

$$\exists l, k_1 \neq l, k_2 \neq l, s_1, s_2 : C_m(a_i^{*k_1,s_1}, l) \geq C_m(a_j^{*k_2,s_2}, l), m = 1..M. \quad (11)$$

5. При отсутствии доминирования, описанного в одном из предыдущих вариантов, можно построить M -мерные выборочные функции распределения критериев для тестирования по одной выборке. Если i -я функция распределения $F_i(C)$ (доминирующая) при любых значениях вектора критериев $C=(C_1, \dots, C_M)$ не больше j -й (доминируемой) $F_j(C)$, а хотя бы в одной точке – строго меньше, то между этими функциями имеет место отношение стохастического доминирования:

$$F_i(C) \leq F_j(C), \forall C; \quad (12)$$

$$\exists C : F_i(C) < F_j(C). \quad (13)$$

Идея стохастического доминирования базируется на предположении о монотонности и независимости всех критериев по предпочтению (чем больше значение критерия, тем лучше, независимо от значений остальных критериев). Стохастическое доминирование (как и доминирование в третьем и четвертом случаях) не дает гарантированного преимущества над доминируемой системой, определяя лишь статистическое преимущество.

Можно выделить следующую иерархию рассмотренных вариантов доминирования. Наиболее явное превосходство i -й системы над j -й наблюдается в первом случае и задается наиболее жесткими ограничениями (4) и (5). Их соблюдение автоматически приводит к соблюдению ограничений (6)–(13) во всех остальных вариантах. При соблюдении ограничений (6) и (7) второго варианта автоматически соблюдаются ограничения (8)–(13) третьего, четвертого и пятого. Третий, четвертый и пятый варианты доминирования связаны зависимостью, однако эта зависимость имеет вероятностный характер и не позволяет по соблюдению одних ограничений сделать вывод о соблюдении других.

Рассмотренные случаи позволяют сделать вывод о той или иной степени превосходства i -й системы над j -й лишь на основе знания о монотонном возрастании критериев по предпочтению инвестора. В других случаях требуется дополнительная информация о важностях критериев для инвестора, дающая возможность построить агрегированный критерий (например, в виде функции полезности). В качестве базовых способов агрегирования критериев можно выделить линейную свертку и свертку Гермейера, которые являются строго возрастающими по предпочтениям инвестора [3, с. 59].

О формирование множества выборок

При изучении стабильности и статистическом сравнении систем управления инвестиционными портфелями важно использовать множество выборок, максимально соответствующее многообразию различных временных рядов рассматриваемых финансовых инструментов. Целесообразно использовать выборки с различными характеристиками волатильности (как по амплитуде, так и по спектру частот). Также интерес представляет исследование на множестве выборок, содержащем как бестрендовые участки, так и в разной степени выраженные тренды, а также моменты смены трендов.

Заключение

При правильно подобранном множестве выборок предлагаемый метод позволяет оценить стабильность качества работы системы управления портфелем при параметрической оптимизации и тестировании на различных выборках, а также сравнить между собой несколько систем. Описанные варианты доминирования позволяют исключить из дальнейшего рассмотрения системы, заведомо уступающие остальным (1-й и 2-й варианты), а также провести более тонкое сравнение остальных систем (3–5-й варианты).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Саркисов В.Г., Саркисов Г.А. Метод оценки и выбора алгоритма принятия инвестиционных решений на основе его параметрической оптимизации // Вестник Самарского государственного технического университета. Сер. Технические науки. – 2009. – № 2(24). – С. 39–46.
2. Пардо Р. Разработка, тестирование, оптимизация торговых систем для биржевого трейдера. – М.: Минакс, 2002. – 224 с.
3. Лотов В.А., Поспелова И.И. Многокритериальные задачи принятия решений. – М.: Макс Пресс, 2008. – 197 с.

Статья поступила в редакцию 20 января 2015 г.

COMPARISON OF INVESTMENT PORTFOLIO MANAGEMENT SYSTEMS AT THE MULTICRITERIA DESCRIPTION OF QUALITY

V.G. Sarkisov, G.A. Sarkisov

Samara State Technical University
244, Molodogvardeyskaya st., Samara, 443100, Russian Federation

The authors continue to research the stability of quality of investment portfolio management systems on the financial market, started in previous paper. The following paper deals with the portfolio management systems optimized within a multi-criteria task. The paper proposes the method of retrospective cross-validation which allows us to estimate the decline in quality caused by the non-optimality of settings. This method also allows us to compare the efficiency of various systems during parametric optimization and testing using multiple series. The paper provides a number of variants of domination of reflections efficient frontiers of set of feasible portfolios for comparison between various optimized systems. The paper also describes the hierarchy of such domination variants.

Keywords: cross-validation, multicriteria optimization, parametric optimization, investment portfolio, efficient frontier.

Vigen G. Sarkisov (Ph.D. (Techn.)), Associate Professor.
Gennady A. Sarkisov (Ph.D. (Techn.)), Associate Professor.