Министерство науки и высшего образования
Российской Федерации
«Московский физико-технический институт
(государственный университет)»
Физтех-школа прикладной математики и информатики
Факультет управления и прикладной математики
Кафедра «Интеллектуальные системы»

Морозов Алексей Олегович

Алгоритмы выбора модели регрессии в больших массивах данных

03.04.01 —Прикладные математика и физика

Выпускная квалификационная работа
(магистерская диссертация)

Научный руководитель:
д. т. н. Моттль Вадим Вячеславович

Москва
2018

Содержание

Аннотация

Обычно в работах по регрессионному анализу необходимость сокращения числа регрессоров объясняется тем, что выбор значений коэффициентов регрессии по всей совокупности наблюдений неизбежно приводит к переобучению и потере обобщающей способности модели. Это означает, что модель, полностью аппроксимирующая массив данных, дает большую ошибку на но-вых объектах, не использовавшихся для оценивания модели. Однако существуют задачи иного рода, в которых достоверно известно, что коэффициенты регрессии отличаются от нуля только в пределах реально существующего ма-лого подмножества большого универсума регрессоров, и поиск этого подмножества является основной целью обработки данных. Именно задачи такого рода рассматриваются в данной работе на примере прикладной задачи восстановления скрытого состава инвестиционного портфеля, представленного временным рядом его периодических доходностей.

Ключевые слова: разряжённая регрессия с ограничениями, предположение о малой коррелированности подмножества активных регрессоров, квадратичное и выпуклое программирование, предположение об равенстве активных коэффициентов регрессии, предположение о равенстве вкладов активных регрессоров в дисперсию целевой переменной, состав инвестиционного портфеля.

1 Введение

В данной работе рассматривается задача регрессионного оценивания при совокупности дополнительных допущений. Предполагается, что коэффициенты регрессии дважды ограничены: отдельным неравенствами неотрицательности вместе с равенством единице их суммы. Кроме того, предполагается, что число регрессоров намного превышает размер выборки, так что даже ограничения неотрицательности и единичной суммы недостаточны для преодоления проблемы переобучения.

Важнейшее дополнительное предположение состоит в том, что коэффициенты регрессии отличаются от нуля только в пределах реально существующего малого подмножества большого универсума регрессоров, и поиск этого подмножества является основной целью обработки данных.

Последнее предположение вытекает из практической задачи восстановления скрытого состава инвестиционного портфеля, представленного временным рядом его периодических доходностей (относительные приращения стоимости портфеля) [1, 2]. Однако нахождение небольшого подмножества фактически активных среди их огромного набора и сильной корреляции, является проблематичным, если нет априорной информации об ожидаемой структуре активного подмножества. Мы рассматриваем три вида априорных предположений, типичных для многих прикладных задач.

Почти всегда уместно предположение, что искомый «истинный» генератор данных предпочитает избегать комбинаций сильно коррелированных регрессоров.

Для задач, связанных с распределением некоторого ресурса по большому числу вариантов вложения, типично предположение, что искомое распределение близко к равномерному на выбранном малом подмножестве вариантов. В задаче восстановления скрытого состава инвестиционного портфеля это означает стремление к диверсификации капиталовложений. Такой принцип априори предполагаемой диверсификации назван здесь Beta Parity, поскольку в финансовой литературе доли капитала принято обозначать греческой буквой Beta.

Более сложное понимание диверсифицированности портфеля заключается не в количественной равномерности распределении капитала, а в выравнивании вкладов разных вложений в общую опасность его потери. Такой принцип чрезвычайно популярен в биржевом бизнесе под названием Risk Parity.

Работа состоит из 9 разделов. В разделе 3 рассмотрена проблема факторного поиска в моделях линейной регрессии при наличии ограничений применительно к задаче восстановления скрытого состава портфеля инвестиционной компании. Показано, что факторный поиск в большом множестве потенциальных коррелированных регрессоров невозможен без принятия некоторых априорных суждений о принципах формирования регрессионной модели, подлежащей оцениванию. В качестве априорной информации предложено использовать предположение о диверсифицированной структуре инвестиционного портфеля, т.е. о распределении капитала по подмножеству активов с разной динамикой колебаний цены. В разделе 4 стремление инвестора к диверсифицированному портфелю математически выражено как идея минимизации дисперсии ожидаемой доходности портфеля, которую обычно упрощенно называют риском потери капитала, поскольку истинный вероятностный риск растет с увеличением этой дисперсии. В разделе 5 идея диверсификации понимается буквально как распределение капитала в одинаковых долях. В разделе 6 формализовано более сложное понимание диверсификации портфеля, чрезвычайно популярное в практике капиталовложений под названием Risk Parity. Раздел 7 содержит результаты модельных экспериментов по восстановлению скрытого состава инвестиционных портфелей. В разделе 8 сформулированы положения, выносимые на за-щиту, а в разделе 9 дана ссылка на опубликованную научную статью.

2 Регуляризация регрессионных моделей данных

Метод наименьших квадратов широко применяется при регрессионном анализе. Пусть $t \in \{1, \dots, T\}$ номера конечного неупорядоченного множества объектов, каждый из которых характеризуется вектором признаков $\mathbf{x}_t = (x_t,i,\ i = 1, \dots, n) \in \mathbb{R}^n$ и целевой переменной $y_t \in \mathbb{R}$. Если предположить, что существует линейная регрессия $y \cong f(\mathbf{x})$, тогда минимизация остаточной суммы квадратов является очевидным и в большинстве случаев отличным средством для нахождения соответствующей оценки коэффициентов регрессии:

$$\begin{cases}\hat{\boldsymbol{\beta}} = \arg\min \sum_{t=1}^{T} (y_t - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})^2 = \arg\min (\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}) \\ \mathbf{y} = (y_1, \ldots, y_T) \in \mathbb{R}^{\mathrm{T}},\ \mathbf{X} = (\mathbf{x}_1 \cdots \mathbf{x}_T)(n \times T),\ \boldsymbol{\beta} = (\beta_i,\ i = 1, \ldots, n) \in \mathbb{R}^n.\end{cases} \tag{1}$$

Понятно, что простое решение соответствующей системы линейных уравнений $\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{X}\mathbf{y}$ дает единственную оценку коэффициентов регрессии только если регрессоры линейно независимы в наборе данных $|\mathbf{X}\mathbf{X}^{\mathrm{T}}| \neq 0$, что возможно только при $n \leqslant T$.

Однако для многих приложений характерно, что число регрессоров значительно превышает размер выборки $n \gg T$. В этом случае неизбежно некоторое ограничение свободы выбора вектора коэффициентов регрессии $\boldsymbol{\beta} \in \mathbb{R}^n$. Существует два принципиально разных способа реализации такого сужения.

Первый способ заключается в добавлении функции регуляризации $V(\mathbf{a})$ критерию МНК (1) для использования некоторых априорных знаний об ожидаемой регрессионной модели:

$$\hat{\boldsymbol{\beta}} = \arg\min \left[V(\boldsymbol{\beta}) + (\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})\right]. \tag{2}$$

Функция регуляризации является штрафом, предназначенным для выражения нежелательности некоторых значений $\boldsymbol{\beta}$ — чем более маловероятным представляется наблюдателю этот вектор коэффициента регрессии, тем больше должен быть штраф $V(\boldsymbol{\beta})$.

Широко известные регуляризации:

- ridge $V(\boldsymbol{\beta}) = \sum_{i=1}^{n} \beta_i^2$ [3, 4]
- bridge $\sum_{i=1}^{n} |\beta_i|^p$, $p > 0$ [5, 6]
- LASSO $\sum_{i=1}^{n} |\beta_i|$ [7]
- Elastic Net $\sum_{i=1}^{n} (\beta_i + \mu|\beta_i|)$ [8]
- SCAD [9]

штрафуют отклонение коэффициентов регрессии от нуля. Фактически, то же самое делает регуляризация на основе некоторой положительно определенной матрицы [10]

$$V(\boldsymbol{\beta}) = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{G}\boldsymbol{\beta}. \tag{3}$$

Главный собственный вектор матрицы регуляризации $\mathbf{G}$ определяет предпочтительную ориентация вектора коэффициентов регрессии в $\mathbb{R}^n$

Однако прикладной задаче анализа инвестиционного портфеля, положившей начало этой работе, оказывается адекватным другой способ регуляризации отличный от штрафной функции (2). Состав портфеля должен оцениваться как доля капитала над набором активов фондового рынка или классов активов, следовательно, коэффициенты регрессии должны в сумме давать единицу

$$\sum_{i=1}^{n} \beta_i = 1, \text{ в векторной форме, } \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{1},\ \mathbf{1} = (1 \cdots 1)^{\mathrm{T}} \in \mathbb{R}^n$$

Кроме того, поскольку мы ограничиваемся рассмотрением только класса взаимных (паевых) инвестиционных фондов, которым разрешено формировать портфели, используя только свой внутренний капитал; в отличие от хедж-фондов, которые могут занимать деньги или активы из внешних источников [11]; что математически приводит к предположению, что коэффициенты регрессии не могут быть отрицательными $\beta_i \geqslant 0$, $i = 1, \ldots, n$, или $\boldsymbol{\beta} \geqslant \mathbf{0}$.

В терминах регуляризирующей функции (2) это ограничение вряд ли поддается обработке:

$$V(\boldsymbol{\beta}) = \begin{cases} 0, & \boldsymbol{\beta} \geqslant \mathbf{0}, \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = 1 \\ \infty, & \text{иначе.} \end{cases}$$

Будет удобнее явно добавить ограничения к критерию МНК (1):

$$\begin{cases} \hat{\boldsymbol{\beta}} = \arg\min(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}) \\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{1},\ \boldsymbol{\beta} \geqslant \mathbf{0} \in \mathbb{R}^n,\ \boldsymbol{\beta} \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^{\mathrm{T}},\ \mathbf{X}(n \times T) \end{cases} \tag{4}$$

3 Проблема факторного поиска. Необходимость регуляризации модели регрессии на основе учета априорной информации о ее скрытой структуре

Обычно в работах по регрессионному анализу необходимость сокращения числа регрессоров объясняется тем, что выбор значений коэффициентов регрессии по всей совокупности наблюдений неизбежно приводит к переобучению и потере обобщающей способности модели. Это означает, что модель, полностью аппроксимирующая массив данных, дает большую ошибку на новых объектах, не использовавшихся для оценивания модели.

Однако существуют прикладные задачи иного рода, в которых достоверно известно, что коэффициенты регрессии отличаются от нуля только в пределах реально существующего малого подмножества большого универсума регрессоров, и поиск этого подмножества является основной целью обработки данных. Именно задачи такого рода рассматриваются в данной работе.

В частности, к таким относится практическая задача восстановления скрытого состава инвестиционного портфеля, представленного временным рядом его периодических доходностей, относительных приращений стоимости портфеля на последовательности периодов владения [2, 12].

Менеджеры инвестиционных компаний, как правило, очень скрытны в отношении того, что они покупают и продают, и долевая структура портфеля $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ обычно скрыта от общественности. Такая информация, будучи восстановленной, представляла бы большой интерес для тех, кто следит за портфелем, в частности, она давала бы инвесторам этого портфеля предварительное оповещение.

Проблема восстановления скрытого распределения капитала в портфеле на основе общедоступных данных была сформулирована Уильямом Шарпом, лауреатом Нобелевской премии 1990 года по экономике [1]. Его метод предполагает анализ временных рядов периодической доходности портфеля, о которых компания обязана отчитываться, совместно с синхронными временными рядами доходности на фондовом рынке активов, предполагаемых к формированию портфеля. Такой принцип восстановления скрытого состава портфеля по публично доступным данным известен под названием Анализ Стиля Инвестиций на Основе Доходности, Returns Based Style Analysis (RBSA) [2]. Эта задача относится к классу задач регрессионного оценивания. Коэффициенты регрессии имеют смысл долевого распределения капитала на множестве биржевых активов, представленных временными рядами их доходностей. Специфика совместного регрессионного анализа массива биржевых данных определяется самой задачей идентификации модели портфеля, заключающейся в поиске небольшого подмножества биржевых активов, в которые действительно вложен капитал, в огромном множестве «всех» биржевых активов.

Коэффициенты регрессии имеют здесь смысл долевого распределения капитала, полностью вложенного в некоторое множество биржевых активов, представленных временными рядами их доходностей, поэтому сумма коэффициентов априори равна единице. В случае паевых фондов (Mutual Funds), которым запрещено заимствовать капитал, возникает также ограничение неотрицательности коэффициентов регрессии:

$$\sum_{i=1}^{n} \beta_i = 1,\ \beta_i \geqslant 0, \text{ в векторной форме, } \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{1},\ \mathbf{1} \in \mathbb{R}^n,\ \boldsymbol{\beta} \geqslant \mathbf{0} \in \mathbb{R}^n \tag{5}$$

К задаче регрессионного анализа приводит модель Шарпа, в которой временной ряд периодических доходностей портфеля $\mathbf{y} = (y_t,\ t = 1,\ldots,T) \in \mathbb{R}^{\mathrm{T}}$ примерно равен линейной комбинации временных рядов доходностей активов (или классов активов) $\boldsymbol{x} = (x_{t,i},\ t = 1,\ldots,T) \in \mathbb{R}^{\mathrm{T}}$, $i = 1,\ldots,n$, с коэффициентами β_i, $i = 1,\ldots,n$, имеющими значение долей капитала, вложенных в каждый из них в начале периода, при условии, что весь бюджет полностью потрачен на инвестиции. Если обозначить все доступное множество регрессоров (активов на рынке ценных бумаг) символом

$$\mathbb{I} = \{i = 1,\ldots,n\},\ n = |\mathbb{I}|, \tag{6}$$

то принцип анализа инвестиций на основе доходности математически выразится как задача квадратичного программирования:

$$\sum_{t=1}^{T} \left(y_t - \sum_{i \in \mathbb{I}} \beta_i x_{t,i} \right)^2 \longrightarrow \min(\beta_i,\ i \in \mathbb{I}),\ \sum_{i=1}^{n} \beta_i = 1,\ \beta_i \geqslant 0,\ i \subset \mathbb{I} \tag{7}$$

в векторной форме,

$$\left(\mathbf{y}-\sum_{i\in\mathbb{I}}\beta_i\boldsymbol{x}_i\right)^{\mathrm{T}}\left(\mathbf{y}-\sum_{i\in\mathbb{I}}\beta_i\boldsymbol{x}_i\right)\longrightarrow\min(\boldsymbol{\beta}\geqslant 0\in\mathbb{R}^n),\ \mathbf{y},\boldsymbol{x}_i\in\mathbb{R}^T,\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta}=\mathbf{1} \qquad (8)$$

В данной работе мы дополнительно наделяем задачу анализа инвестиций требованием найти малое подмножество активов, в которое фактически вложен капитал исследуемой инвестиционной компании. В математических терминах это означает поиск в (6), (7) подмножества регрессоров, коэффициенты при которых отличны от нуля $\hat{\mathbb{I}}=\{i:\ \beta_i>0\}\in\mathbb{I}=\{i=1,\ldots,n\}$. Такую задачу мы будем называть задачей факторного поиска (Factor Search).

Еще раз отметим, что целью поиска подмножества активных регрессоров является не улучшение обобщающей способности модели, а определение фактически действующих факторов $\boldsymbol{x}_i\in\mathbb{R}^T$, $i\in\hat{\mathbb{I}}\subset\mathbb{I}$, сформировавших анализируемый сигнал $\mathbf{y}\in\mathbb{R}^T$.

4 Регуляризация на основе предположения о стремлении инвестора минимизировать риск потери капитала

4.1 Принцип минимизации риска

Пусть $\mathbf{X}(n\times T)$ — полная заданная совокупность биржевых активов, в развернутом виде

$$\begin{pmatrix} x_{11} & x_{21} & \ldots & x_{Tn} & \\ x_{12} & x_{22} & \ldots & x_{Tn} & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \ldots & x_{Tn} & \end{pmatrix}(n\times T),\ \mathbf{X}=(\mathbf{x}_1\cdots\mathbf{x}_n)=\begin{pmatrix}\mathbf{x}_1^{\mathrm{T}}\\ \vdots \\ \mathbf{x}_n^{\mathrm{T}}\end{pmatrix},\mathbf{x}_t\in\mathbb{R}^n,\ \boldsymbol{x_t}\in\mathbb{R}^T. \qquad (9)$$

Естественные ограничения равенства и неравенства могут оказаться недостаточными для наделения критерия регрессионной оценки (4) свойством фактор-поиска.

Это будет явно видно, если мы запишем (4) в полу-скалярной форме:

$$\begin{cases} (\mathbf{y}-\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y}-\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}) = \left(\mathbf{y}-\sum i=1^n \beta_i \boldsymbol{x}_i\right)^{\mathrm{T}}\left(\mathbf{y}-\sum i=1^n \beta_i \boldsymbol{x}_i\right) \longrightarrow \min(\beta_1,\ldots,\beta_n) \\ \sum\limits_{i=1}^{n} \beta_i = 1,\ \beta_i \geqslant 0,\ i=1,\ldots,n \end{cases} \tag{10}$$

Основным препятствием является корреляция между регрессорами $\boldsymbol{x} = (x_{1,i}\cdots x_{T,i}) \in \mathbb{R}^T$, $i = 1,\ldots,n$, что делает едва различимыми различные комбинации $[\boldsymbol{x}_i,\ i \in \hat{\mathbb{I}} \subset \mathbb{I} = \{1,\ldots,n\}]$.

Эта трудность может быть существенно ослаблена, если априори предположить, что регрессоры в искомой комбинации слабо коррелируют между собой. Чем меньше корреляция между случайными переменными $(x_i, i = 1,\ldots,n)$, тем меньше дисперсия их линейной комбинации $y = \sum_{i=1}^n \beta_i x_i$ с фиксированной нормой вектора коэффициентов $\sum_{i=1}^n \beta_i^2$. Таким образом, если мы хотим найти подмножество слабо коррелированных регрессоров, мы должны минимизировать квадратичную форму $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta} \to \min$, $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta} = \mathrm{const}$.

4.2 Математический критерий построение минимально рискованного портфеля на заданном множестве биржевых активов

Мы исходим из предположения, что портфель, временной ряд доходности $(y_t, t = 1,\ldots,T)$ которого анализируется, был составлен командой его менеджеров с целью получения большей доходности при приемлемом уровне риска.

Пусть доходности активов в последующие периоды времени $\mathbf{x}_t = (x_{t,1}\cdots x_{t,n})$, $t = 1,\ldots,T$ рассматривается как совокупность реализаций случайного вектора, имеющего математическое ожидание $\bar{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\} \in \mathbb{R}^n$ и ковариационную матрицу $\boldsymbol{\Sigma} = \mathbb{E}\{(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^{\mathrm{T}}\}(n \times n)$. Здесь мы опускаем вопрос о том, являются ли доходности в последовательность $\mathbf{x}_t, t = 1,\ldots,T$ зависимыми или нет, предполагается, что соответствующий временной ряд стационарен, и рассматривается его константная доходность. Тогда, доходность портфеля $y_t = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_t$ также следует рассматривать как стационарный случайный процесс, и его доходность на заданном отрезке времени будет случайной величиной с математическим ожиданием $\bar{y} = \boldsymbol{\beta}^{\mathrm{T}}\bar{\mathbf{x}}$, а именно, ожидаемой прибылью (доходностью), и дисперсией $\sigma^2 = \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}$.

Чем больше дисперсия доходности, тем больше вероятность низкой и даже отрицатель-ной доходности, и тем хуже становится портфель. Стремление инвестора к диверсифицированному портфелю математически выражено как идея минимизации дисперсии ожида-емой доходности портфеля, которую обычно упрощенно называют риском потери капи-тала, поскольку истинный вероятностный риск растет с увеличением этой дисперсии.

Средние значения доходностей активов $\bar{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\} \in \mathbb{R}^n$ малы по сравнению с их колебаниями вокруг среднего значения. Поэтому вместо ковариационной матрицы $\boldsymbol{\Sigma} = \mathbb{E}\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathrm{T}}\}$ достаточно использовать просто матрицу скалярных произведений временных рядов доходностей

$$\boldsymbol{\Sigma} = \mathbb{E}\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathrm{T}}\} \cong \frac{1}{T}\sum t = 1^T \mathbf{x}_t \mathbf{x}_t^{\mathrm{T}} = \frac{1}{T}\mathbf{X}\mathbf{X}^{\mathrm{T}} \tag{11}$$

в обозначениях (1).

Таким образом, мы приходим к следующей задаче формирования портфеля в заданном полном множестве активов:

$$\begin{cases} J(\boldsymbol{\beta}) = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta} \to \min(\boldsymbol{\beta} \in \mathbb{R}^n) \\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = 1,\ \boldsymbol{\beta} \geqslant 0 \end{cases} \quad \begin{cases} D(\boldsymbol{\beta}) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\beta_i \to \min(\beta_i) \\ \sum\limits_{i=1}^{n}\beta_i = 1,\ \beta_i \geqslant 0,\ i = 1,\ldots,n, \end{cases} \tag{12}$$

Этот критерий выражает стремление инвестора минимизировать риск потери капитала.

Это задача квадратичного программирования, решаемая стандартными вычислительными методами.

Правда, вычислительная сложность такой задачи весьма высока, она является кубичной относительно полного числа активов n, которая может составлять много сотен и даже тысячи. Проблемы снижения вычислительной сложности очень актуальна, однако здесь мы ограничиваемся решением только концептуальных вопросов идентификации скрытого состава инвестиционного портфеля.

Следует ожидать, что в результате минимизации критерия (12) $\hat{\boldsymbol{\beta}} = \arg\min J(\boldsymbol{\beta})$ будет определен состав наименее рискованного портфеля, включающего лишь некоторое число активов с β-коэффициентами, отличными от нуля $\beta_i > 0$:

$$\hat{\boldsymbol{\beta}} \in \mathbb{R}^n,\ \hat{\mathbb{I}} = \{i : \beta_i > 0\} \subset \mathbb{I} = \{i = 1, \ldots, n\},\ \hat{n} = |\hat{\mathbb{I}}| < n.$$

Опыт показывает, что критерий (12) очень селективен на реальных множествах активов, временные ряды доходностей которых $\{\boldsymbol{x}_i \in \mathbb{R}^T,\ i = 1, \ldots, n\}$ (9) являются, как правило, очень сильно коррелированными. В разделе (7) мы увидим, на массивах реальных временных рядов месячных доходностей активов критерий (12) дает вырожденный результат, подавляя большинство активов и оставляя лишь несколько самых стабильных по доходности, демонстрирующих очень малое ее колебание вокруг среднего значения.

Однако жестокая реальность заключается в том, что всегда средняя доходность биржевого актива тем меньше, чем меньше дисперсия ее колебаний. Как следствие, портфель, построенный по принципу минимального риска (12), оказывается малодоходным. Кто не рискует, тот не пьет шампанского!

Для преодоления этого противоречия в работе [13] предлагается использовать принцип формирования портфеля, предложенный Гарри Марковицем [14], и заключающийся выборе портфеля, обеспечивающего выбранный инвестором баланс между взаимно противоречивыми требованиями доходности и малой рискованности. Это баланс выражается так называемым коэффициентом Risk Tolerance, являющимся индивидуальной психологической характеристикой инвестора.

4.3 Управление селективностью формирования портфеля по принципу минимальной рискованности

Итак, критерий (12), требующий выбрать портфель с минимальным риском потери капитала приводит к вырожденному решению вложить весь капитал в один самый малорискованный актив либо в очень небольшое число таких активов. В силу непреодолимых законов биржевого равновесия эти активы оказываются и самыми принцип Risk Parity малодоходными.

Но даже если инвестор согласен на малую доходность ради минимизации риска, то надежность такого портфеля является кажущейся. Суждение о малом риске основано на предположении, что ковариационная матрица (11) известна. Если она вдруг изменится, например, по причинам, внешним по отношению к естественному биржевому равновесию [15], то именно этот актив может стать очень рискованным, что приведет к фатальной рискованности простейшего портфеля. Иначе говоря, прямое следование критерию минимизации риска входит в противоречие с требованием диверсификации портфеля, рассмотренным ниже в Разделах (5) и (6).

Построить портфель с бóльшим числом активов можно, только допустив увеличение риска.

Минимизация наивного критерия (12) дает минимальное возможное значение риска к данной множестве активов:

$$d_{\min} = \min_{\mathbf{1}^{\mathrm{T}}\boldsymbol{\beta}=1, \boldsymbol{\beta}\geqslant 0} \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta},\ \hat{\boldsymbol{\beta}}_{d_{\min}} = \big(\operatorname*{arg\,min}_{\mathbf{1}^{\mathrm{T}}\boldsymbol{\beta}=1, \boldsymbol{\beta}\geqslant 0} \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}\big) \in \mathbb{R}^n \tag{13}$$

$$\hat{\mathbb{I}}_{d_{\min}} = \{i : \hat{\beta}_{d_{\min}} > 0\} \subset \mathbb{I} = \{i = 1, \ldots, n\},\ \hat{n}_{d_{\min}} = |\hat{\mathbb{I}}_{d_{\min}}| \tag{14}$$

есть состав минимально рискованного портфеля.

С другой стороны, очевидно, что

$$\begin{aligned} &d_{\max} = \max_{i=1,\ldots,n} (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_i) &&\text{— дисперсия доходности самого волатильного актива}\\ &i_{\max} \in \mathbb{I} = \{i = 1, \ldots, n\} &&\text{— его номер}\\ &\hat{\mathbb{I}}_{\max} = \{i : i = i_{\max}\} &&\text{— множество из одного самого «плохого» актива.} \end{aligned} \tag{15}$$

Совокупность всех чисел

$$d_{\min} \leqslant d \leqslant d_{\max} \tag{16}$$

определяет однопараметрическое семейство портфелей, состав и β-коэффициенты каждого из которых является решением задачи оптимизации

$$\begin{aligned} &\hat{\boldsymbol{\beta}}_d = (\arg\min \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}) \in \mathbb{R}^n\\ &\mathbf{1}^{\mathrm{T}}\boldsymbol{\beta} = 1,\ \boldsymbol{\beta} \geqslant 0,\ \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta} = d, \end{aligned} \quad \hat{\mathbb{I}}_d = \{i : i = \beta_{d,i} > 0\}. \tag{17}$$

Для каждого значения d (16) это задача квадратичной оптимизации при одном квадратичном ограничении, в отличие от задачи квадратичного программирования в случае только линейных ограничений. Мы не будем рассматривать здесь способы решения такой, вообще говоря, классической задачи, хотя ее исключительно простая структура позволит построить простой и быстрый алгоритм решения.

Для нас важно, что совокупность значений параметра d определяет однопараметрическое семейство всех подмножеств активов

$$\mathbb{J} = \{\hat{\mathbb{I}}_d : d_{\min} \leqslant d \leqslant d_{\max}\} \tag{18}$$

которые образуют семейство портфелей, оптимальных по критерию минимума риска в некотором подмножестве активов. Крайние значения $d = d_{\min}$ и $d = d_{\max}$ вряд ли приемлемы. Но никакие другие подмножества из 2^n подмножеств в множестве всех активов $\hat{\mathbb{I}} \subset \mathbb{I} = \{i = 1, \ldots, n\}$ не могут образовать оптимальный портфель.

Мы видим прямую аналогию с регуляризацией восстановления скрытого состава портфеля в пределах однопараметрического семейства портфелей, эффективных по Марковицу [13]. Даже параметр регуляризации имеет сходный смысл, назначая приемлемое значение риска. Будем называть параметр d (16) параметром риска.

4.4 Определение состава инвестиционного портфеля в предположении, что он построен по принципу минимума риска

Пусть наблюдается некоторый портфель над множеством активов, представленных их периодическими доходностями $\{\boldsymbol{x}_i \in \mathbb{R}^T,\ i = 1, \ldots, n\}$ (9), представленный временным рядом его доходностей $\mathbf{y} = (y_t,\ t = 1, \ldots, T) \in \mathbb{R}^T$, но β-коэффициенты $\boldsymbol{\beta}^* = (\beta_i^* \geqslant 0,\ i = 1, \ldots, n) \in \mathbb{R}^n$ неизвестны. Согласно модели Шарпа

$$\mathbf{y} \cong \sum_{i=1}^{n} \beta_i^* \boldsymbol{x}_i. \tag{19}$$

Наблюдатель хочет узнать значения β-коэффициентов, но при $n \gg T$ это сделать, вообще говоря, невозможно (раздел 3). Единственное, что можно с уверенность утверждать, это то, что портфель построен на небольшом неизвестном множестве активов:

$$\mathbb{I}^* \subset \mathbb{I} = \{i = 1, \ldots, n\},\ \beta_i^* > 0,\ i \in \mathbb{I}^*,\ \sum_{i \in \mathbb{I}^*} \beta_i^* = 1. \tag{20}$$

Это равносильно утверждению, что портфель построен на всем множестве активов, но доли вложения капитала положительны лишь на части активов:

$$\beta_i^* \geqslant 0,\ i \in \mathbb{I} = \{i = 1, \ldots, n\},\ \beta_i^* > 0, i \in \mathbb{I}^*,\ \beta_i^* = 0, i \notin \mathbb{I}^*,\ \sum_{i \in \mathbb{I}^*} \beta_i^* = 1. \tag{21}$$

Предположим, что портфель построен по принципу минимума риска (12). Тогда можно считать известным, что портфель построен на подмножестве активов, входящем а однопараметрическое множество (18). Остается лишь выбрать значение параметра риска (16).

Опыт показывает, что параметрическое семейство (18) содержит только подмножества, размер которых не превышает числа наблюдений $n_d < T$.

Для каждого конкретного значения параметра d скрытые значения β-коэффициентов $\boldsymbol{\beta}_d^* \in \mathbb{R}^{n_d}$ естественно оценивать по критерию наименьших квадратов (1) в пределах предполагаемого подмножества активов (регрессоров). Опыт показывает, что параметрическое семейство (18) содержит, как правило, только подмножества, размер которых меньше числа наблюдений. На случай, если это не так, введем в критерий небольшую дополнительную ридж-регуляризацию с малым коэффициентом $\alpha \to 0$. Эвристически отбросим ограничения $\sum_{i \in \mathbb{I}_d} \beta_i^2 = 1$ и $\beta_i \geqslant 0,\ i \in \mathbb{I}_d$:

$$\hat{\boldsymbol{\beta}}_d = (\beta_{d,i},\ i \in \hat{\mathbb{I}}_d) = \arg\min\{\alpha \sum_{i \in \mathbb{I}_d} \beta_i^2 + \sum_{t=1}^{T} \left(y_t - \sum_{i \in \mathbb{I}_d} \beta_i x_{t,i} \right)^2 \}. \tag{22}$$

Очевидно, что средняя остаточная сумма квадратов$\sum_{t=1}^{T} \left(y_t - \sum_{i \in \mathbb{I}_d} \hat{\beta}_i x_{t,i} \right)^2$ будет уменьшаться с увеличением числа регрессоров в активном подмножестве $n = |\hat{\mathbb{I}}_d|$, и наиболее подходящее подмножество нельзя найти по минимуму $S_{\hat{\mathbb{I}}_d}$. Измерить степень соответствия подмножества $\hat{\mathbb{I}}_d$ массиву данных

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix} \in \mathbb{R}^T,\ \boldsymbol{x}_i = (x_{1,i} \cdots x_{T,i})^{\mathrm{T}} \in \mathbb{R}^T,\ i = 1,\ldots,n,\ \mathbf{X} = (\mathbf{x}_1 \cdots \mathbf{x}_T) = \begin{pmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{x}_1^{\mathrm{T}} \end{pmatrix} (n \times T) \tag{23}$$

позволяет критерий Leave One Out (LOO).

Критерий LOO заключается в следующем. Удалим из множества наблюдений $\mathbb{T} = \{s = 1,\ldots,T\}$ одно наблюдение с номером t и рассмотрим сокращенный массив данных

$$\begin{aligned} &\mathbb{T}^{(t)} = \{s = 1,\ldots,t-1,t+1,\ldots,T\}, \\ &\boldsymbol{x}_i^{(t)} = (x_{1,i} \cdots x_{t-1,i} \cdots x_{T,i})^{\mathrm{T}} \in \mathbb{R}^{T-1},\ i = 1,\ldots,n \end{aligned} \tag{24}$$

Решим задачу, аналогичную (22), T раз, всякий раз удалив одно наблюдение t:

$$\hat{\boldsymbol{\beta}}_d^{(t)} = (\beta_{d,i}^{(t)},\ i \in \hat{\mathbb{I}}_d) = \arg\min\{\alpha \sum_{i \in \mathbb{I}_d} \beta_i^2 + \sum_{t \in \mathbb{T}^{(t)}} \left(y_t - \sum_{i \in \mathbb{I}_d} \beta_i x_{t,i} \right)^2 \}. \tag{25}$$

Средняя остаточная сумма квадратов по полному множеству наблюдений и есть критерий Leave-One-Out:

$$LOO(d) = \frac{1}{T} \sum_{t \in \mathbb{T}} \left(y_t - \sum_{i=1}^{n} \beta_{d,i}^{(t)} x_{t,i} \right)^2. \tag{26}$$

Остается найти оценку истинного состава портфеля как точку минимума функции $LOO(d)$:

$$\mathbb{I}^* = \arg\min LOO(d). \tag{27}$$

Для этого не нужно перебирать пробные значения параметра d на сетке. Опыт показывает, что эта функция, как правило, унимодальна, и точку ее минимума можно найти методом золотого сечения в интервале $d_{\min} \leqslant d \leqslant d_{\max}$ (15)-(16), вычислив при этом значения $LOO(d)$ лишь для нескольких значений параметра.

Прямое вычисление каждого значения функции $LOO(d)$ согласно (26) имеет высокую вычислительную сложность. К счастью, это не требуется. Существует очень простой способ вычисления критерия LOO, всего один раз решив, казалось бы, бесполезную задачу наименьших квадратов для конкретного значения параметра d по всем наблюдениям (22) [16]. Соответствующая формула имеет вид

$$LOO(d) = \frac{1}{T} \sum_{t \in \mathbb{T}} \left(\frac{y_t - \sum_{i \in \mathbb{I}_d} \beta_{\mathbb{I}_d,i} x_{t,i}}{1 - \mathbf{x}_{\mathbb{I}_d,t}^{\mathrm{T}} \left(\mathbf{X}_{\mathbb{I}_d} \mathbf{X}_{\mathbb{I}_d}^{\mathrm{T}} + \frac{1}{\alpha} \mathbf{I}_{\mathbb{I}_d} \right)^{-1} \mathbf{x}_{\mathbb{I}_d,t}} \right)^2 \tag{28}$$

где согласно (24) $\mathbf{x}_{\mathbb{I}_d,t} \in \mathbb{R}^{n_d}$, а матрица $(\mathbf{X}_{\mathbb{I}_d} \mathbf{X}_{\mathbb{I}_d}^{\mathrm{T}})$ имеет размер $(n_d \times n_d)$.

5 Регуляризация на основе предположения о равномерном распределении капитала с целью количественной диверсификации портфеля

5.1 Принцип Beta Parity

Под диверсификацией понимают принцип распределение инвестируемых денежных капиталов между как можно более различными объектами вложений с целью снижения риска возможных потерь капитала или доходов от него.

С математической точки зрения выбор инвестиционного портфеля понимается как, во-первых, выбор в множестве «всех» активов, доступных на биржевом рынке $\mathbb{I} = \{i = 1, \ldots, n\}$, $n = |\mathbb{I}|$, некоторого подмножества для вложения капитала $\hat{\mathbb{I}} \subset \mathbb{I}$, $\hat{n} < n$, и, во-вторых выбор долевого распределения капитала на этом подмножестве β_i, $i \in \mathbb{I}$,

$$\boldsymbol{\beta} = (\beta_i,\, i \in \mathbb{I}),\ \beta_i > 0, i \in \hat{\mathbb{I}},\ \beta_i = 0, i \notin \hat{\mathbb{I}},\ \sum_{i \in \hat{\mathbb{I}}} \beta_i = \sum_{i=1}^{n} \beta_i = 1. \tag{29}$$

В этих терминах диверсифицированный портфель выражается вектором β-коэффициентов, удовлетворяющим условию

$$\beta_i = \beta_j, \quad i,j \in \hat{\mathbb{I}}, \quad \beta_i = 0,\, i \notin \hat{\mathbb{I}} \tag{30}$$

Такой простейший вид диверсификации портфеля будем называть принципом

Beta Parity:

$$
\begin{cases}
\sum\limits_{i=1}^{n}\begin{pmatrix} 2\mu\beta_i, & \beta_i \leqslant \mu \\ \mu^2+\beta_i^2, & \beta_i > \mu \end{pmatrix} + c\sum\limits_{t=1}^{T}\left(y_t - \sum\limits_{i=1}^{n}\beta_i x_{t,i}\right)^2 \longrightarrow \min(\beta_1,\ldots,\beta_n) \\
\sum\limits_{i=1}^{n}\beta_i = 1,\ \beta_i \geqslant 0,\ i = 1,\ldots,n
\end{cases}
\tag{31}
$$

5.2 Построение портфеля Beta Parity на заданном множестве биржевых активов

Формирование модельного портфеля производится при естественном ограничении на дисперсию всего портфеля:

$$
d_{\min} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j), \quad d_{\max} = \max_i\{\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_i\}. \tag{32}
$$

Чтобы доли каждого актива были одинаковые, необходимо решить следующую оптимизационную задачу при ограничении на дисперсию портфеля:

$$
\begin{cases}
\sum\limits_{i=1}^{n}\beta_i^2 \longrightarrow \min(\beta_1,\ldots,\beta_n) \\
\sum\limits_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\beta_i = d,\ \sum\limits_{i=1}^{n}\beta_i = 1,\ \beta_i \geqslant 0,\ i = 1,\ldots,n
\end{cases}
\tag{33}
$$

При увеличении параметра d от $d_{\min}$ к $d_{\max}$ возрастает селективность и часть регрессоров перестает быть активной.

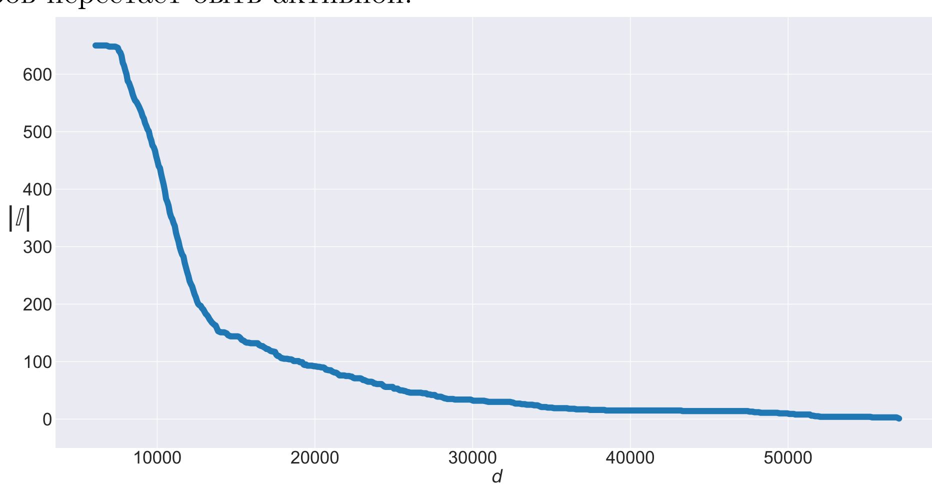

Рис. 1: Размер портфеля Beta Parity при увеличении параметра d

6 Предположение о диверсификации портфеля путем равномерного распределения вкладов активов в риск потери капитала

6.1 Принцип Risk Parity

Доходность портфеля и ее дисперсия

$$y = \sum_{i=1}^{n} \beta_i \boldsymbol{x}_i, \quad D(\boldsymbol{\beta}) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\beta_i = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\right)\beta_i}_{D_i(\boldsymbol{\beta})} = \sum_{i=1}^{n} D_i(\boldsymbol{\beta}) \quad (34)$$

Здесь $D_i(\boldsymbol{\beta})$ — вклад i-го актива в дисперсию доходности портфеля $D(\boldsymbol{\beta})$.

Условие Risk Parity:

$$D_i(\boldsymbol{\beta}) = D_j(\boldsymbol{\beta}) \text{ для всех регрессоров } i,j = 1,\ldots,n \quad (35)$$

Критерий Risk Parity — сочетание требований малого риска и диверсификации:

$$\begin{cases} \sum\limits_{i=1}^{n}\left[nD_i - \sum\limits_{j=1}^{n} D_j\right]^2 \longrightarrow \min(\beta_1,\ldots,\beta_n) \\ D_i = \left(\sum\limits_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\right)\beta_i,\ \beta_i \geqslant 0,\ i = 1,\ldots,n,\ \sum\limits_{i=1}^{n}\beta_i = 1 \end{cases} \quad (36)$$

После решения задачи (36) получим портфель с минимальной дисперсией сформированный по принципу Risk Parity. Однако все активы войдут в него.

Идея управления размером портфеля Risk Parity — параметр размера $d_{\min} \leqslant d \leqslant d_{\max}$, $d_{\max} = \max\limits_i(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_i)$

Параметрическое семейство портфелей Risk Parity

$$\begin{cases} \sum\limits_{i=1}^{n} D_i^2 \longrightarrow \min(\beta_1,\ldots,\beta_n) \\ \sum\limits_{i=1}^{n} D_i = d,\ D_i = \left(\sum\limits_{j=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j)\beta_j\right)\beta_i \\ \sum\limits_{i=1}^{n}\beta_i = 1,\ \beta_i \geqslant 0,\ i = 1,\ldots,n \end{cases} \quad \Longrightarrow \quad \begin{array}{c} \underbrace{\mathbb{I} = \{i = 1,\ldots,n\}}_{\text{все активы}} \\ \underbrace{\mathbb{I}_d = \{i : \hat{\beta}_i > 0\} \subseteq \mathbb{I}}_{\text{портфель с параметром } d} \end{array} \quad (37)$$

Параметр d — психологическая характеристика инвестора. Чем меньше d – тем меньше риск, но больше размер портфеля. Чем больше d – тем больше риск, но меньше портфель. Изменяя параметр d.

Получаем однопараметрическое семейство моделей Risk Parity:

$$\mathbb{J} = \{\hat{\mathbb{I}}_d : \ d_{\min} \leqslant d \leqslant d_{\max}\} \tag{38}$$

Наилучшую модель $d_{\min} \leqslant d \leqslant d_{\max}$ из семейства $\mathbb{J}$, которая наилучшим образом соответствует временному ряду $\mathbf{y} \in \mathbb{R}^T$, получим по критерию Leave-One-Out:

$$LOO(d) = \frac{1}{T}\sum_{t\in\mathbb{T}}\left(y_t - \sum_{i=1}^{n} \beta_{d,i}^{(t)} x_{t,i}\right)^2 \tag{39}$$

Прямое вычисление каждого значения функции $LOO(d)$ согласно (39) имеет высокую вычислительную сложность. К счастью, это не требуется. Существует очень простой способ вычисления критерия LOO. Соответствующая формула имеет вид

$$LOO(d) = \frac{1}{T}\sum_{t\in\mathbb{T}}\left(\frac{y_t - \sum_{i\in\mathbb{I}_d} \beta_{\mathbb{I}_d,i} x_{t,i}}{1 - \mathbf{x}_{\mathbb{I}_d,t}^{\mathrm{T}}\left(\mathbf{X}_{\mathbb{I}_d}\mathbf{X}_{\mathbb{I}_d}^{\mathrm{T}}\right)^{-1}\mathbf{x}_{\mathbb{I}_d,t}}\right)^2 \tag{40}$$

согласно обозначениям (24).

6.2 Построение портфеля Risk Parity на заданном множестве биржевых активов

Были произведены анализ всех биржевых активов и формирование модельного портфеля для каждого значения параметра d (16).

При любом множестве биржевых активов первым делом следует вычислить диапазон всех допустимых значениях d, при которых критерий будет иметь смысл, а именно $d_{\min}$ и $d_{\max}$.

При $d_{\min}$ все регрессоры активны и мы имеет портфель Risk Parity на всём множестве биржевых активов. Такой портфель имеет наименьшую дисперсию при наименьшей средней доходности.

При $d_{\max}$ активным остаётся только один регрессов и портфель Risk Parity со-

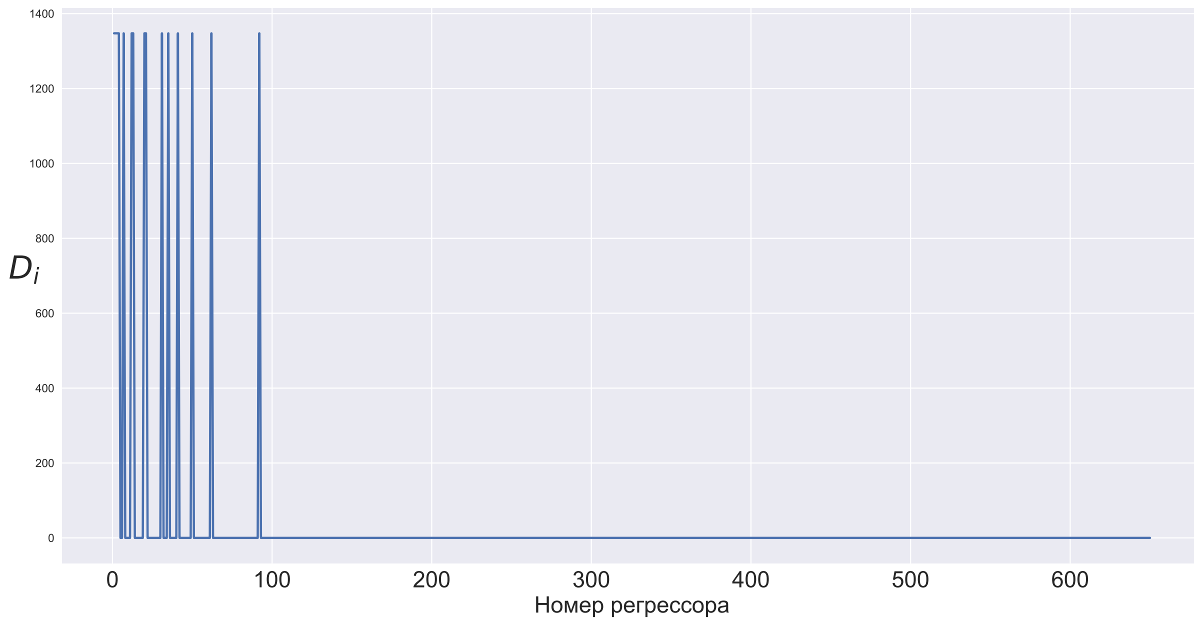

Рис. 2: Равенство коэффициентов D_i при 16 активных регрессорах

стоит из одного биржевого актива. Этот портфель с наибольшей дисперсией при наибольшей средней доходности.

$$d_{\max} = \max_i \{\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_i\} \tag{41}$$

Таблица 1: Размер портфеля

d	$d_{\min} = 1024$	3785	5252	6029	6202	6288	9568	13883	18198	22513
$\lvert\mathbb{I}_d\rvert$	650	650	649	648	647	646	382	112	37	22
d	26828	31143	35458	39774	44089	48404	52719	56862	56948	$d_{\max} = 57034$
$\lvert\mathbb{I}_d\rvert$	16	11	9	8	7	7	4	3	2	1

Можно сказать, что размер портфеля монотонно убывает, за исключением некоторых вырожденных промежутков. Это возникает из-за того, что регрессоры сильно коррелированы. Критерий очень селективен при увеличении параметра риска d, оставляет мало биржевых активов.

Как можно видеть из таблицы (2), портфели не являются вложенными друг в друга. Есть регрессоры, которые присутствуют начиная с какого-то момента, остальные же могут как появляться, так и пропадать в определенный момент.

Таблица 2: Состав портфеля

Размер портфеля	Номера индексов
1	**7**
2	**3**, 7
3	**2**, 3, 7
4	**1**, 2, 3, 7
5	1, 2, 3, 7, **35**
6	1, 2, 3, 7, **34**, 35
7	1, 2, 3, 7, **12**, **13**, 35
8	1, 2, 3, 7, 12, 13, **16**, 35
8	1, 2, 3, **4**, 7, **20**, **21**, **50**
9	1, 2, 3, 4, 7, 20, 21, **31**, 50
10	1, 2, 3, 4, 7, 20, 21, 31, **41**, 50

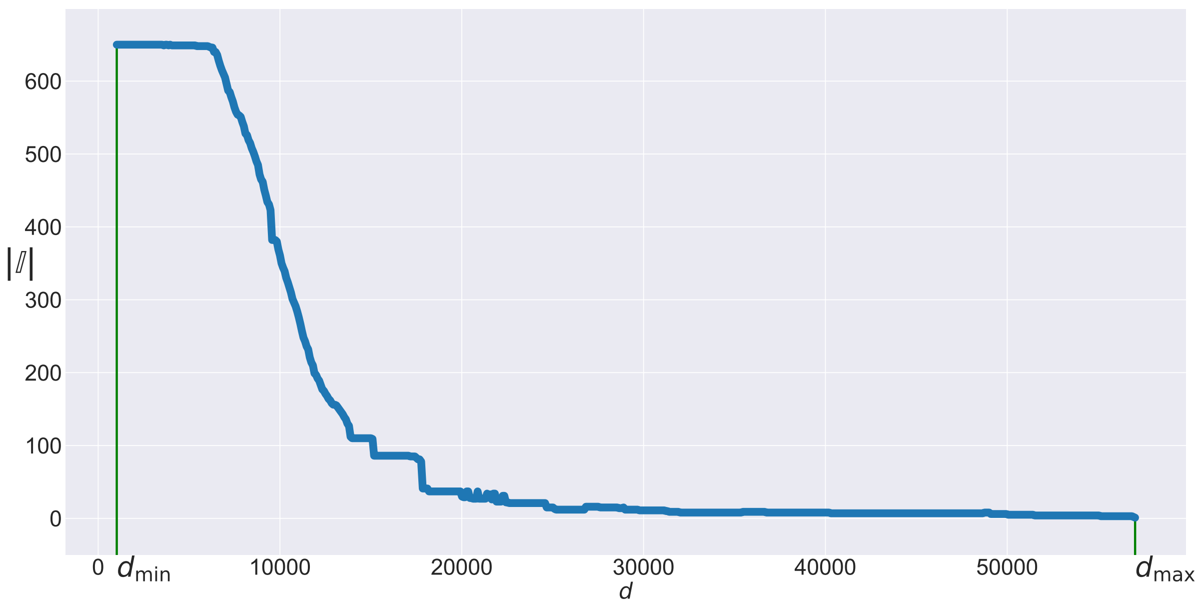

Рис. 3: Размер портфеля при увеличении параметра d

7 Эксперименты по восстановлению скрытого состава инвестиционного портфеля

Размерность матрицы временных доходностей (9) составляет ($650 = n \times T = 251$). Всего было выбрано $n = 650$ значений параметра риска d равномерно на сетке от $d_{\min}$ до $d_{\max}$.

В разделе (6.2) было посчитано, что

$$d_{\min} = 1024, \quad d_{\max} = 57034 \tag{42}$$

Для каждого $l = 1, \ldots, n$ был построен временной ряд

$$\mathbf{y}^l = \sum_{i=1}^{n} \beta_i^l \boldsymbol{x}_i + \xi \mathbf{1},\, \mathbf{1} \in \mathbb{R}^T \tag{43}$$

где ξ — 10 % нормальный белый шум с нулевым средним и дисперсией

$$\sigma_\xi^2 = 0.1 \frac{1}{T} \left(\sum_{i=1}^{n} \beta_i^l \boldsymbol{x}_i \right)^{\mathrm{T}} \left(\sum_{i=1}^{n} \beta_i^l \boldsymbol{x}_i \right) \tag{44}$$

Для восстановления состава инвестиционного портфеля был выбран временной ряд, сформированный при некотором значении риска d^*.

Применяем LOO по всем (рассматриваемым) значениям $d_{\min} \leqslant d \leqslant d_{\max}$, фактически по всем возможным портфелям.

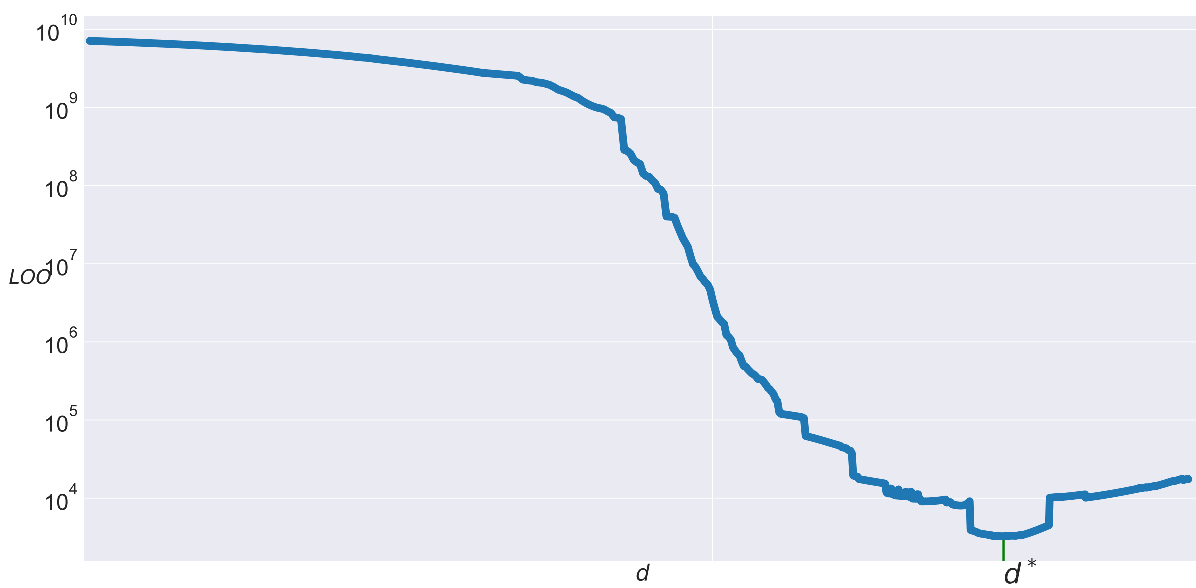

Рис. 4: $LOO(d)$ по составу портфеля $\hat{\mathbb{I}}_d$

Находим минимум $LOO(d)$, например, методом золотого сечения.

8 Положения, выносимые на защиту

1) Математическая формулировка задачи регуляризации регрессионной модели инве-стиционного портфеля по принципу диверсификации
2) Методология использования в качестве априорной информации предположения о примерном равенстве активных коэффициентов регрессии – Beta Parity.
3) Выпуклый критерий Beta Parity для управляемого селективного отбора регрессоров Modulus Quadratic и алгоритм его минимизации.
4) Методология использования в качестве априорной информации предположения о равенстве вкладов активов в общий риск портфеля – Risk Parity.
5) Выпуклый критерий Risk Parity для управляемого селективного отбора регрессоров и алгоритм его минимизации.
6) Концепция выбора модели предъявленного портфеля в однопараметрических се-мействах Beta Parity и Risk Parity.
7) Быстро-вычисляемый критерий LOO

9 Публикация по материалам работы

O. Krasotkina, M. Markov, V. Mottl, D. Babichev, I. Pugach, A. Morozov. Constrained Regularized Regression Model Search in Large Sets of Regressors. *Machine Learning and Data Mining in Pattern Recognition. Lecture Notes in Computer Science*, Springer, 2018 (to appear).

Список литературы

[1] *Sharpe, W. F.* Asset allocation: Management style and performance measurement / W. F. Sharpe // *Journal of portfolio Management.* — 1992. — Vol. 18, no. 2. — Pp. 7–19.

[2] *Markov, M.* Dynamic style analysis and applications / M. Markov, V. Mottl, I. Muchnik. — 2004.

[3] *Horel, A.* Application of ridge analysis to regression problems / A. Horel // *Chemical Engineering Progress.* — 1962. — Vol. 58. — Pp. 54–59.

[4] *Vinod, H. D.* Recent advances in regression methods / H. D. Vinod, A. Ullah. — Marcel Dekker Incorporated, 1981. — Vol. 41.

[5] *Frank, L. E.* A statistical view of some chemometrics regression tools / L. E. Frank, J. H. Friedman // *Technometrics.* — 1993. — Vol. 35, no. 2. — Pp. 109–135.

[6] *Fu, W. J.* Penalized regressions: the bridge versus the lasso / W. J. Fu // *Journal of computational and graphical statistics.* — 1998. — Vol. 7, no. 3. — Pp. 397–416.

[7] *Tibshirani, R.* Regression shrinkage and selection via the lasso / R. Tibshirani // *Journal of the Royal Statistical Society. Series B (Methodological).* — 1996. — Pp. 267–288.

[8] *Zou, H.* Regularization and variable selection via the elastic net / H. Zou, T. Hastie // *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* — 2005. — Vol. 67, no. 2. — Pp. 301–320.

[9] *Fan, J.* Variable selection via nonconcave penalized likelihood and its oracle properties / J. Fan, R. Li // *Journal of the American statistical Association.* — 2001. — Vol. 96, no. 456. — Pp. 1348–1360.

[10] *Kakade, S. M.* Regularization techniques for learning with matrices / S. M. Kakade, S. Shalev-Shwartz, A. Tewari // *Journal of Machine Learning Research.* — 2012. — Vol. 13, no. Jun. — Pp. 1865–1890.

[11] *Tyson, E.* Investing for Dummies / E. Tyson. — John Wiley & Sons, 2011.

[12] *Markov, M.* Dynamic analysis of hedge funds // The 3rd IASTED International Conference on Financial Engineering and Applications, ACTA Press, Cambridge. — 2006.

[13] *Ilya, P.* — Identifying the model of the investment portfolio by the time series of its returns and a large array of returns on exchange-traded assets. — Masterthesis, MIPT, 2018.

[14] *Markowitz, H.* Portfolio selection / H. Markowitz // *The journal of finance.* — 1952. — Vol. 7, no. 1. — Pp. 77–91.

[15] *Kaletsky, A.* How mr soros made a billion by betting against the pound' / A. Kaletsky // *The Times, 26th October.* — 1992.

[16] *Cawley, G. C.* Fast exact leave-one-out cross-validation of sparse least-squares support vector machines / G. C. Cawley, N. L. Talbot // *Neural networks.* — 2004. — Vol. 17, no. 10. — Pp. 1467–1475.