Московский государственный университет имени М. В. Ломоносова

Факультет вычислительной математики и кибернетики

Кафедра математических методов прогнозирования

Бабичев Дмитрий Олегович

Регуляризация регрессионной модели доходности инвестиционного портфеля для учета априорной информации о его структуре

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель:

академик РАН

Журавлев Юрий Иванович

Научные консультанты:

д.т.н. *Моттль В.В.*,

к.ф-м.н. *Красоткина О.В.*

Москва, 2018

Содержание

1 Введение

1.1 Доходность инвестиционного портфеля

Цель инвестиционной компании – вложить имеющийся капитал в активы таким образом, чтобы максмизировать прибыль и при этом минимизировать риск разорения. Пусть имеется n биржевых активов, β_i – доля капитала, вложенная в i-ый актив. Вектор коэфициентов долевого распределения капитала по выбранному множеству активов, которые принято называть бета-коэфициентами $\boldsymbol{\beta} = \big(\beta_1, ..., \beta_n\big)$, полностью определяет состав инвестиционного портфеля компании. Естественными ограничениями для портфеля являются неотрицательность бета коэффициентов $\beta_i \geq 0$ и суммирование их в единицу $\sum_{i=1}^{n} \beta_i = 1$. Ограничения вида $\beta_i \geq 0$ связаны с тем фактом, что активы могут быть приобретены только с использованием внутреннего капитала без заимствования денег или активов из внешних источников. Большинству инвестиционных фондов заимствования запрещены. Инвестиционные компании, занимающие отрицательные позиции $\beta_i < 0$, называются хедж-фондами. В этой работе мы не рассматриваем хедж-фонды.

Как правило, структура инфестиционного портфеля скрывается от общественности. Такая информация, будучи востановленной, представляла бы большой интерес для тех, кто следит за портфелем. Проблема восстановления скрытого распределения капитала в портфеле из общедоступных данных была сформулирована Уильямом Шарпом, лауреатом Нобелевской премии по экономике 1990 года [7]. 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.

Поскольку фактический портфель может содержать сотни и даже тысячи биржевых инструментов, Шарп предложил приближать итоговую доходность портфеля небольшим числом рыночных индексов [8], представляющих определенные классы активов и инвестиционные стили. Этот подход известен под названием анализа стиля инветиций компании на основе сравнения временых рядов доходностей (returns) ее портфеля с временными рядями некоторой совокупности биржевых индексов (Return Based Style Analysis, RBSA).

Периодическая доходность – это процент прибыли от инвестиций в финансовый инструмент за определенный период времени. Пусть $z_{\text{beg},i}$ и $z_{\text{end},i}$ – стоимости актива i в начале периода владения и в конце. Отношение $x_i = \frac{z_{\text{end},i}-z_{\text{beg},i}}{z_{\text{beg},i}}$ характеризует относительный прирост стоимости (доходность) актива. Аналогично, отношение $y = \frac{z_{\text{end},p}-z_{\text{beg},p}}{z_{\text{beg},p}}$ выражает относительный прирост стоимости (доходность) портфеля за данный период владения.

Последовательность доходностей за различные периоды владения образуют временные ряды y_t и $x_{t,i},\ i = 1, ..., n$. Эти последовательности считаются известными, потому что компании обязаны периодически публиковать свои доходности, а стоимости активов публикуются онлайн или в печатных финансовых СМИ.

В модели Шарпа периодическая доходность портфеля равна линейной комбинации периодических доходностей активов или классов активов с коэффициентами, имеющими значение долей капитала, вложенных в каждый из них в начале каждого периода владения, при условии, что весь бюджет полностью потрачен на инвестиции:

$$\sum_{t=1}^{T}\left(y_t - \sum_{i=1}^{n}\beta_i x_{t,i}\right)^2 \to \min, \quad \sum_{i=1}^{n}\beta_i = 1,\ \beta_i \geq 0,\ i = 1, ..., n \tag{1}$$

Задача (1) относится к классу квадратичных задач с линейными ограничениями.

1.2 Factor search

Множество всевозможных активов $\boldsymbol{I} = \{1, ..., n\}$ обычно достаточно велико, поэтому несмотря на ограничения, почти всегда существует портфель $\boldsymbol{\beta}$, обеспечивающий нулевой остаток в критерии (1), но не отражающий реальный состав анализируемого портфеля. Поэтому обычно рассматривают более простую задачу – поиск активного подмножества $\hat{\boldsymbol{I}} \subset \boldsymbol{I}$, такого, что $\beta_i > 0\ \forall i \in \hat{\boldsymbol{I}}$, $\beta_i = 0\ \forall i \notin \hat{\boldsymbol{I}}$. Такая формулировка подразумевает своего рода регуляризацию, существенно уменьшающую степень свободы переменных. Однако встает вопрос: как находить это активное подмножество? В данной работе предлагается учесть, помимо естественных ограничений, дополнительные априорные предположения о рациональности портфеля с точки зрения его полезности как финансового инструмента.

2 Априорные предположения о составе анализируемого портфеля

Будем исходить из того, что портфель $(y_t,\ t = 1, ..., T)$ был составлен командой менеджеров с целью получения наибольшей доходности при приемлемом уровне риска.

Будем рассматривать временные ряды доходностей активов $\boldsymbol{x}_t = (x_{t,1}, ..., x_{t,n}),\ t = 1, ..., T$, как реализации векторной случайной величины с математическим одижанием $\bar{\boldsymbol{x}} = \mathbb{E}(\boldsymbol{x}) \in \mathbb{R}^n$ и ковариационной матрицей $\boldsymbol{\Sigma} = \mathbb{E}\big[(\boldsymbol{x}-\bar{\boldsymbol{x}})(\boldsymbol{x}-\bar{\boldsymbol{x}})^T\big](n\times n)$. Опуская вопрос о независимости значений последовательности $\boldsymbol{x}_t$ во времени, будем считать, что она стационарна. Тогда доходность портфеля $y_t = \boldsymbol{\beta}^T\boldsymbol{x}_t$ может рассматриваться как стационарный случайный процесс, а его сечение – как векторная случайная величина с математическим ожиданием $\bar{y} = \boldsymbol{\beta}^T\bar{\boldsymbol{x}}$, выражающим ожидаемую среднюю прибыль, и дисперсией $\sigma^2 = \boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta}$, которую обычно понимают как степень риска убытка.

С одной стороны, чем больше ожидаемая доходность, тем лучше. Но с другой стороны, повышение доходности принципиально возможно только за счет увеличения ее дисперсии, и следовательно, увеличения риска убытка. Чем больше дисперсия доходности, тем больше вероятность низкой и даже отрицательной доходности, и тем хуже портфель.

Теория построения оптимального портфеля была разработана Гарри Марковицем в 1952 году [3] и явилась составной частью мотивации присуждения ему вместе с Уильямом Шарпом Нобелевской премии по экономике в 1990 году. Согласно этой теории только тот портфель считается рациональным, который имеет минимальную дисперсию при фиксированной ожидаемой доходности ($\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta} \to \min,\ \bar{\boldsymbol{x}}^T\boldsymbol{\beta} = \text{const}$), и наоборот – гарантирует максимальную ожидаемую доходность при фиксированной дисперсии ($\bar{\boldsymbol{x}}^T\boldsymbol{\beta} \to \max,\ \boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta} = \text{const}$). Легко показать, что оба требования покрываются условием

$$U(\boldsymbol{\beta}, \mu) = \mu\bar{\boldsymbol{x}}^T\boldsymbol{\beta} - \boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta} \to \max,\ \boldsymbol{\beta} \geq \mathbf{0},\ \mathbf{1}^T\boldsymbol{\beta} = 1. \tag{2}$$

Квадратичную функцию $U(\boldsymbol{\beta}, \mu)$ называют Portfolio Utility [6], т.е. полезностью портфеля. Параметр μ, называемый Risk Tolerance, интерпретируют как индиивидуальную готовность инвестора к риску убытка в надежде получить высокий доход, а обратную величину $\frac{1}{\mu}$, называемую Risk Aversion, понимают как количественное выражение обратной позиции инвестора, его неприятие риска.

Таким образом, при фиксированном математичесом ожидании $\bar{\boldsymbol{x}}$ и ковариационной матрице $\boldsymbol{\Sigma}$ доходностей активов существует однопараметрическое семейство портфелей, в котором каждый портфель не может быть улучшен сразу по двум взаимно противоречивым показателям. Оптимальность по Марковицу полностью эквивалентна оптимальности по Парето, и может рассматриваться как частный случай последней в контексте портфельной задачи.

Именно этот факт, обеспечивающий само существование равновесия на фондовом рынке в сообществе инвесторов, не склонных к риску, привел к модели ценообразования биржевых активов (Capital Asset Pricing Model, CAPM) [6]. Портфели, не отвечающие условию (2), не могут быть признаны приемлемыми с точки зрения теории Марковица, однако невозможно судить о том, какое значение толерантности к риску лучше. Это внутренняя психологическая характеристика каждого конкретного инвестора.

В настоящей работе предлагается использовать класс функций полезности (2) в качестве параметрического семейства регуляризационных функций для задачи выбора модели исследуемого инвестиционного портфеля, которую мы называем здесь Factor Search:

$$V(\boldsymbol{\beta}, \mu) = \boldsymbol{\beta}^T \boldsymbol{\Sigma} \boldsymbol{\beta} - \mu \bar{\boldsymbol{x}}^T \boldsymbol{\beta} \to \min. \tag{3}$$

Здесь терпимость к риску μ является гиперпараметром, а статистики $\bar{\boldsymbol{x}}$ и $\boldsymbol{\Sigma}$ нужно оценивать по историческим данным как выборочное среднее и выборочную ковариационную матрицу. Однако если предполагать, что эти характеристики относительно стабильны в период, по которому мы оцениваем портфель, то не будет большой ошибкой оценивание их по данным за наблюдаемый период $1, ..., T$:

$$\bar{\boldsymbol{x}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_t, \;\; \boldsymbol{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} (\boldsymbol{x}_t - \bar{\boldsymbol{x}})(\boldsymbol{x}_t - \bar{\boldsymbol{x}})^T. \tag{4}$$

3 Регуляризация для отбора регрессоров

3.1 Центральная регуляризация

Как упоминалось ранее, естественных ограничений в критерии (1) может оказаться недостаточно для поиска активного подмножества активов. Основным препятствием является корреляция между отдельными регрессорами $\boldsymbol{x}_i = (x_{1,i}, ..., x_{T,i}) \in \mathbb{R}^T,\ i = 1, ..., n$, которая делает трудноразличимыми различные комбинации регрессоров $[\boldsymbol{x}_i,\ i \in \hat{\boldsymbol{I}} \subset \boldsymbol{I} = \{1, 2, ..., n\}]$. Можно существенно ослабить влияние корреляции на отбор регрессоров, если предполагать, что для создания портфеля отбирались некоррелированные активы, что способствует снижению риска. Чем меньше корреляция между случайными величинами $(x_i,\ i = 1, ..., n)$, 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 функцию $\boldsymbol{\beta}^T \boldsymbol{\Sigma} \boldsymbol{\beta} \to \min,\ \boldsymbol{\beta}^T \boldsymbol{\beta} = \text{const}$.

Это только одно из возможных априорных предположений, в общем случае любая положительно определенная матрица $\boldsymbol{G}$ может выражать интуицию пользователя, она даже не обязана зависеть от данных. Учитывая все это, наша задача будет выглядеть так:

$$\begin{cases} \boldsymbol{\beta}^T \boldsymbol{G} \boldsymbol{\beta} + c(\boldsymbol{y} - \boldsymbol{X}^T \boldsymbol{\beta})^T (\boldsymbol{y} - \boldsymbol{X}^T \boldsymbol{\beta}) \to \min_{\beta_1, ..., \beta_n} \\ \boldsymbol{1}^T \boldsymbol{\beta} = 1, \boldsymbol{\beta} \geq \boldsymbol{0} \end{cases} \tag{5}$$

Матрица квадратичной функции регуляризации указывает на априорное предпочтительное направление в пространстве регрессоров. Точкой минимума регуляризации является нулевой вектор $\boldsymbol{\beta} = \boldsymbol{0}$ – центр пространства. Поэтому такую регуляризацию будем называть центральной.

3.2 Смещенная регуляризация

Однако может случиться так, что априорные знания предполагают не только предпочтительное направление вектора коэффициентов, но и их конкретные значения. В частности, сама идея отбора регрессоров (Factor Search) предполагает, что число активных регрессоров меньше их общего числа n. Это означает, что мы хотим заставить процедуру регрессионного анализа отдавать предпочтение нулевым значениям коэффициентов регрессии $\beta_i = 0$, оставляя активными, т.е. положительными $\beta_i > 0$, только те из них, которые наиболее необходимы для аппроксимации анализируемого временного ряда $y_t \cong \sum_{i=1}^{n} \beta_i x_{t,i}$. Это может быть сделано путем выбора следующей квадратичной функции:

$$(\boldsymbol{\beta} - \boldsymbol{\beta}_0)^T \boldsymbol{G} (\boldsymbol{\beta} - \boldsymbol{\beta}_0) \to \min, \ \boldsymbol{\beta}_0 < \mathbf{0}, \tag{6}$$

в которой смещение ожидаемого вектора бета-коэффициентов $\boldsymbol{\beta}_0$ в отрицательный квадрант противоречит предположению о неотрицательности его компонент $\boldsymbol{\beta} \geqslant \mathbf{0}$, что и приводит к обнулению "лишних"коэффициентов. Такую регуляризацию будем называть смещенной. Заметим, что если $\boldsymbol{\beta}_0 = -\mu \boldsymbol{I}$, то регляризация примет вид $\boldsymbol{\beta}^T \boldsymbol{G} \boldsymbol{\beta} + \mu \mathbf{1}^T \boldsymbol{G} \boldsymbol{\beta} \to \min$.

Такой вид смещеной регуляризации мы уже использовали, когда говорили о теории оптимального портфеля Г. Марковица и У. Шарпа (3).

Также легко увидеть некоторую аналогию с Elastic Net $\sum_{i=1}^{n} (\beta_i^2 + \mu |\beta_i|)$. Если принять $\boldsymbol{G} = \boldsymbol{I}$, то регуляризация примет вид $\boldsymbol{\beta}^T \boldsymbol{\beta} + \mu \mathbf{1}^T \boldsymbol{\beta} \to \min$, который с учетом ограничений неотрицательности $\boldsymbol{\beta} \geqslant \mathbf{0}$ полностью соответствуют Elastic Net.

4 Вычислительные эксперименты

4.1 Исходные данные и условия экспериментов

Целью экспериментов является изучение возможности процедуры, описанной в разделах 2 и 3, оценивать фактическую разреженность модели регрессии для большого числа числа коррелированных переменных. В качестве набора регрессоров мы использовали 650 временных рядов месячных доходностей биожевых индексов $\boldsymbol{x}_t = (x_{t,i}, i = 1, ..., n) \in \mathbb{R}^n, \ n = 650$, за 20 лет, $T = 240$:

$$\boldsymbol{X} = (\boldsymbol{x} \in \mathbb{R}^n, \ t = 1, ..., T)(n \times T).$$

Если упорядочить собственные значения ковариационной матрицы $\boldsymbol{\Sigma}$, то можно заметить резкое падение собственных значений: $\lambda_1 = 16369.8, \ \lambda_2 = 1435.2, \ \lambda_3 = 1076.2, \ \lambda_{20} = 134.0 < 0.01\lambda_1, \ \lambda_{50} = 32.2 < 0.002\lambda_1$, что свидетельствует об очень существенной кореляции между нидексами. Идея экспериментов состоит в том, чтобы попытаться восстановить набор бета-фоэффициентов, составляющих наблюдаемый инвестиционный портфель, заведомо разумный с точки зрения теории Марковица, но с неизвестным значением рискованности Risk Tolerance.

Мы построили пять портфелей по критерию

$$\begin{cases} (\beta^*_{\mu,1}, ..., \beta^*_{\mu,n}) = \arg\min\big(\boldsymbol{\beta^T}\boldsymbol{\Sigma}\boldsymbol{\beta} - \mu\bar{\boldsymbol{x}}^T\boldsymbol{\beta}\big), \\ \mathbf{1}^T\boldsymbol{\beta} = 1,\ \boldsymbol{\beta} \geq \mathbf{0}, \end{cases}$$

для разных значений Risk Tolerance μ:

$$\mu_1 = 0,\ \mu_2 = 10,\ \mu_3 = 30,\ \mu_4 = 200,\ \mu_5 = 1000.$$

Как и ожидалось, полученные портфели оказались разреженными, то есть в каждом из пяти потфелей оказалось небольшое число положительных коэффициентов по сравнению с общим числом биржевых индексов $n = 650$:

Risk Tolerance	$\mu_1 = 0$	$\mu_2 = 10$	$\mu_3 = 30$	$\mu_4 = 200$	$\mu_5 = 1000$
Number of active indexes	$m_1 = 13$	$m_2 = 11$	$m_3 = 10$	$m_4 = 7$	$m_5 = 1$

Для каждого набора коэффициентов $(\beta^*_{\mu,i},\ i \in \boldsymbol{I}^*_\mu)$ мы вычислили временной ряд доходности

$$\boldsymbol{y}_\mu = (y_{\mu,t},\ t = 1,..,T) \in \mathbb{R}^T,\ y_{\mu,t} = \sum_{i\in\boldsymbol{I}^*_\mu} \beta^*_{\mu,i} x_{t,i} + \xi_t,$$

как последовательность случайных величин с 10%-ой дисперсией шума

$$\sigma^2(\xi_t) = 0.1\frac{1}{T}\sum_{t=1}^{T}\left(y_t - \sum_{i\in\boldsymbol{I}^*_\mu}\beta^*_{\mu,i}x_{t,i}\right)^2.$$

4.2 Поиск факторов без регуляризации

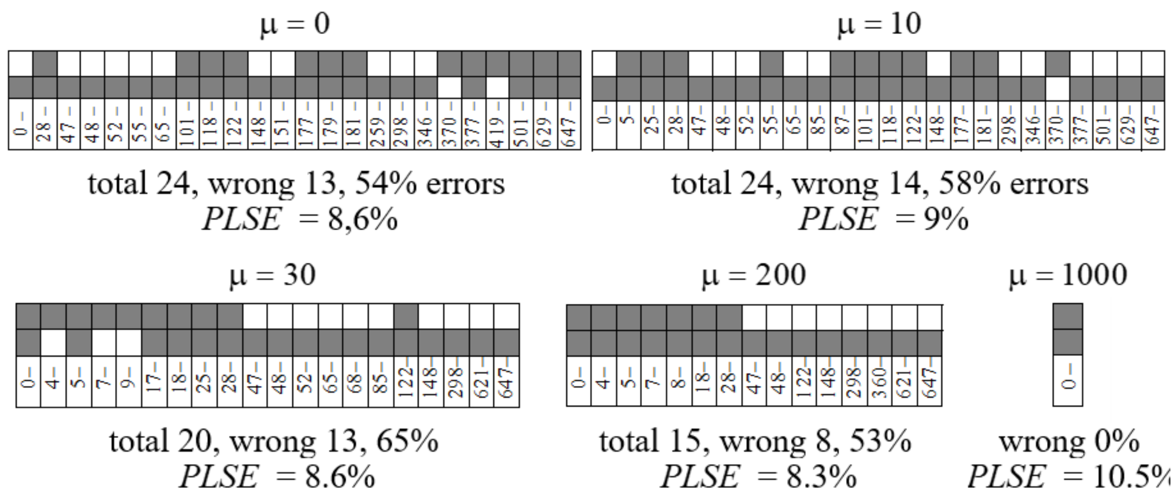
Для начала убедимся, что без регуляризации найти активное подмножество регрессоров невозможно. Находить коэффициенты будем, просто минимизируя квадрат ошибки с ограничениями:

$$\begin{cases} (\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{\beta})^T(\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{\beta}) \to \min(\boldsymbol{\beta}), \\ \mathbf{1}^T\boldsymbol{\beta} = 1,\ \boldsymbol{\beta} \geq \mathbf{0}. \end{cases} \tag{7}$$

После нахождения вектора коэффициентов $\hat{\boldsymbol{\beta}}_\mu$, вычислим относительную ошибку на настощем активном подмножестве Percentage Least Squares Error (PLSE):

$$\mathrm{PLSE} = \frac{\sum_{t=1}^{T} \left(y_t - \sum_{i \in \bar{\boldsymbol{I}}_\mu} \widehat{\beta_{i,\mu}} x_{t,i} \right)^2}{\sum_{t=1}^{T} \left(\sum_{i \in \bar{\boldsymbol{I}}_\mu} \widehat{\beta_{i,\mu}} x_{t,i} \right)^2} \times 100. \tag{8}$$

В следующей таблице представлены результаты оценивания состава портфеля без регуляризации по критерию (7). Белые квадраты означают отсутствие соответствующего регрессора, а серые – присутствие. В верхней строке закрашены квадраты, соответствующим настоящим регрессорам, а в нижней – оцененная.



Из рисунка видно, что за исключением тривиального случая большого значения Risk Tolerance, более 50% регрессоров ошибочно идентифицируются, ято говорит о необходимости регуляризации на основе предположении о принадлежности портфеля к классу оптимальных по Марковицу.

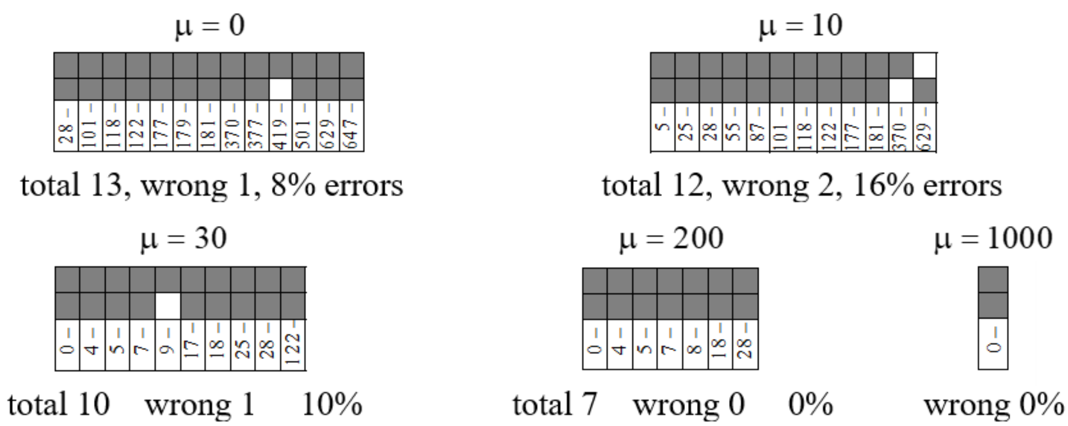
4.3 После введения регуляризации

Как показано выше, простая минимизация квадрата ошибки не позволяет оценить активное подмножество. Теперь попробуем минимизировать критерий с регуляризацией по Марковицу (3):

$$\begin{cases} \boldsymbol{\beta}^T \boldsymbol{\Sigma} \boldsymbol{\beta} - \mu \bar{\boldsymbol{x}}^T \boldsymbol{\beta} + \dfrac{1}{T} (\boldsymbol{y}_\mu - \boldsymbol{X}^T \boldsymbol{\beta})^T (\boldsymbol{y}_\mu - \boldsymbol{X}^T \boldsymbol{\beta}) \to \min(\boldsymbol{\beta}) \\ \mathbf{1}^T \boldsymbol{\beta} = 1, \ \boldsymbol{\beta} \geq \mathbf{0}. \end{cases} \tag{9}$$

Параметр μ будем подбирать с помощью критерия Leave-one-out (LOO).

Результаты:



5 Заключение

В ходе работы над ВКР были изучены некоторые аспекты регрессионного анализа в больших наборах регрессоров при двух видах ограничений – неотрицательности коэффициентов регрессии и суммирование их в единицу. Дополнительное предположение состоит в том, что коэффициенты регрессии отличаются от нуля только в пределах реально существующего малого подмножества большой совокупности сильно коррелированных регрессоров. Основной целью работы был поиск факторов, а именно нахождение этого подмножества. Еще одна трудность заключается в том, что найти небольшое активное подмножество среди огромного набора коррелированных факторов весьма проблематично, если не имеется некоторой априорной информации об ожидаемой структуре портфеля. Было предложено два способа введения дополнительной регуляризации в факторном поиске: центральная регуляризация, квадратичная матрица которой должна указывать на некоторое априорно предпочтительное направление в размерном пространстве регрессоров и смещенная регуляризация через штрафование числа активных регрессоров.

5.1 Актуальные вопросы

1. В теории построения портфелей помимо теории Марковица существуют другие критерии, в частности – Risk Parity, суть которого в том, чтобы уравнять вклады в общий риск портфеля отдельных активов. Можно ли с помощью регуляризации восстанавливать портфели такой структуры?

2. В данной работе предполагалось, что состав портфеля не менялся в течение оцениваемого периода, что может не соответствовать действительности.

3. Также, в данной работе не были затронуты вопросы вычислительной сложности. Мы ограничились всего 650 регрессорами, однако на реальных биржах количество активов может достигать десятки тысяч.

Список литературы

[1] *Bro R., De Jong S.* A fast non-negativity-constrained least squares algorithm // *Journal of chemometrics.* — 1997. — Vol. 11, no. 5. — Pp. 393–401.

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

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

[4] *Meinshausen N. et al.* Sign-constrained least squares estimation for high-dimensional regression // *Electronic Journal of Statistics.* — 2013. — Vol. 7. — Pp. 1607–1631.

[5] Methods of hyperparameter estimation in time-varying regression models with application to dynamic style analysis of investment portfolios / O. Krasotkina, V. Mottl, M. Markov et al. // International Conference on Machine Learning and Data Mining in Pattern Recognition / Springer. — 2017. — Pp. 431–450.

[6] *Sharpe W. F.* Capital asset prices: A theory of market equilibrium under conditions of risk // *The journal of finance.* — 1964. — Vol. 19, no. 3. — Pp. 425–442.

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

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