УДК 004.021; 519.863

Алгоритмическая реализация методологии оценивания состава инвестиционных портфелей

А.О. Морозов[1], В.В Моттль[2]

[1]Московский физико-технический институт (государственный университет)

[2] Федеральный исследовательский центр «Информатика и управление» Российской Академии Наук

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.

$$\begin{cases} \sum_{t=1}^{T}\left(y_t - \sum_{i=1}^{n}\beta_i x_{t,i}\right)^2 \to \min(\beta_1,\ldots,\beta_n) \\ \sum_{i=1}^{n}\beta_i = 1, \beta_i \geq 0, i = 1,\ldots,n \quad \text{задача квадратичного программирования} \end{cases}$$

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 [3], поскольку в финансовой литературе доли капитала принято обозначать греческой буквой Beta.

Регрессионный критерий Beta Parity.

Параметр селективности $\mu \geq 0$

$$\begin{cases} \sum_{i=1}^{n}\begin{pmatrix} 2\mu\beta_i, & \beta_i \leq \mu \\ \mu^2 + \beta_i^2, & \beta_i > \mu \end{pmatrix} + c\sum_{t=1}^{T}\left(y_t - \sum_{i=1}^{n}\beta_i x_{t,i}\right)^2 \to \min(\beta_1,\ldots,\beta_n) \\ \sum_{i=1}^{n}\beta_i = 1, \beta_i \geq 0, i = 1,\ldots,n \quad \text{задача выпуклого программирования} \end{cases}$$

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. Регрессионный критерий имеет следующий вид:

$$\left\{\begin{array}{l} \sum_{i=1}^{n} D_i^2 \to \min(\beta_i, i \in \mathbb{I}),\\ \sum_{i=1}^{n} D_i = d,\ D_i = \left(\sum_{i=1}^{n} (\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j)\beta_j \right)\beta_i,\\ \sum_{i=1}^{n} \beta_i = 1, \beta_i \geq 0, i \in \mathbb{I}. \end{array}\right\}$$

$$\underbrace{\mathbb{I} = \{i = 1, \dots, n\}}_{\text{все активы}}$$

$$\underbrace{\mathbb{I}_d = \{i : \hat{\beta}_i > 0\} \subseteq \mathbb{I}}_{\text{портфель с параметром d}}$$

Меньше d – меньше риск, больше портфель.

Больше d – больше риск, меньше портфель

Параметр d – психологическая характеристика инвестора

С вычислительной точки зрения задача восстановления скрытого состава инвестиционного портфеля по наблюдаемому временному ряду его периодических доходностей (Returns Based Style Analysis) есть задача регрессионного анализа при совокупности дополнительных допущений. Во-первых, коэффициенты регрессии (искомое долевое распределение капитала) ограничены неравенствами неотрицательности вместе с равенством единице их суммы. Во-вторых, предполагается, что число регрессоров (число возможных биржевых активов) намного превышает размер выборки. Дополнительное предположение состоит в том, что коэффициенты регрессии отличаются от нуля только в пределах реально существующего малого подмножества регрессоров (искомого состава портфеля) в многократно большем множестве всех активов. В качестве неизбежной проблемно-ориентированной регуляризации мы используем естественное предположение, что анализируемый портфель рационально построен его администрацией в соответствии с теорией эффективных портфелей Гарри Марковица. Мы обеспечиваем линейную вычислительную сложность алгоритма поиска состава портфеля в очень большом множестве всех биржевых активов, в то время, как сложность по относительно небольшому числу наблюдений остается полиномиальной.

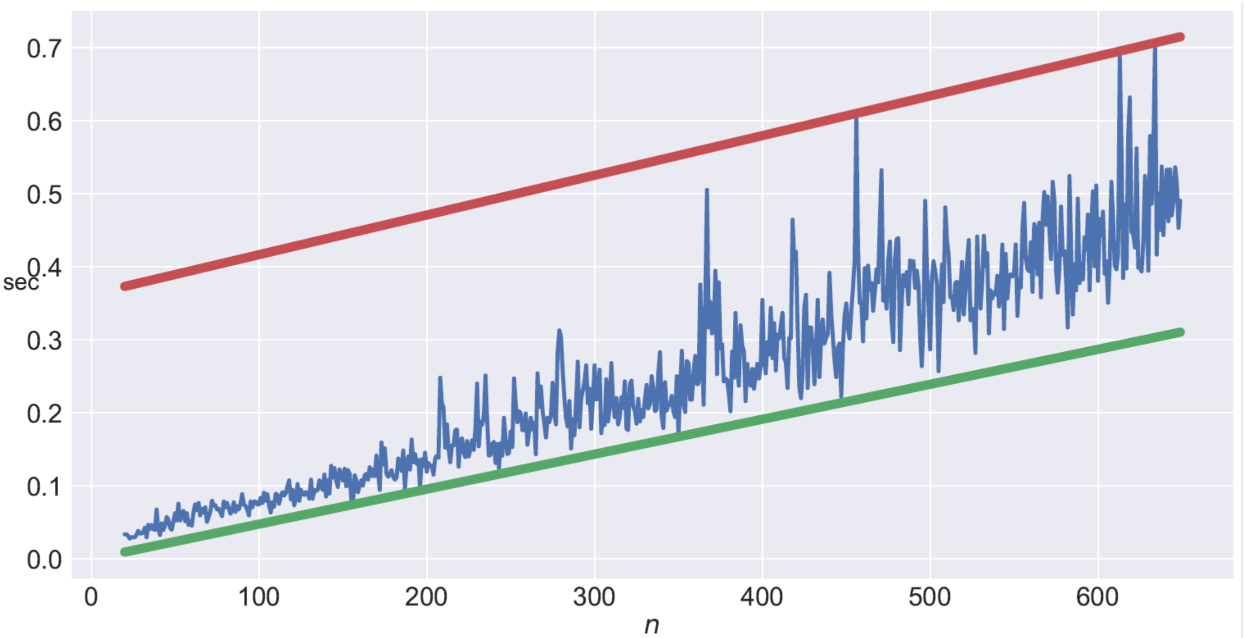

Рис. 1. Линейная вычислительная сложность по количеству активов

Литература

1. Sharpe, W.F. Asset allocation: Management style and performance measurement. The Journal of Portfolio Management, Winter 1992, pp. 7-19.
2. Markov M., Mottl V., Muchnik I. Dynamic Style Analysis and Applications. SSRN Electronic Journal,·August 2004.
3. O. Krasotkina, M. Markov, V. Mottl, D. Babichev, I. Pugach, A. Morozov. Constrained Regularized Regression Model Search in Large Sets of Regressors. Lecture Notes in Artificial Intelligence, Vol. 1035, Springer, 2018, pp. 1-15.