Применение рекуррентного обучения с подкреплением для решения задачи портфельной оптимизации

Титов Владислав Валерьевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: vladislav.v.titov@yandex.ru

Процессы на финансовых рынках являются очень популярным объектом для исследований, которые нередко порождают теории и методы, применяющиеся в самых разных сферах жизни.

Существует множество задач, связанных с моделированием процессов на финансовых рынках. Их постановки различаются по частоте совершения сделок, количеству используемых биржевых инструментов, информации, используемой для принятия решений. В данной работе внимание сконцентрировано на задаче оптимизации портфеля финансовых инструментов, которая заключается в оптимальном распределении капитала между активами, входящими в портфель.

Одним из наиболее актуальных подходов к решению задачи портфельной оптимизации является применение методов рекуррентного обучения с подкреплением (RRL) [3]. Применение RRL подразумевает построение дифференцируемой параметрической модели на основании оптимизации некоторого критерия, то есть минимизации или максимизации выбранного функционала (например, при помощи метода градиентного спуска).

Пусть R_t — доходность торгового инструмента, $t \in [0, T]$. В качестве оптимизируемого функционала будем использовать коэффициент Шарпа, который, помимо доходности, учитывает рисковость построенного портфеля и формально задается следующим образом:

$$S_T = \frac{\overline{R_T}}{\sigma(R_T)},$$

где $\overline{R_T} = \frac{1}{T}\sum\limits_{k=0}^{T} R_k$, $\sigma(R_T) = \sqrt{\frac{1}{T}\sum\limits_{k=0}^{T}(R_k - \overline{R_T})^2}$

В работе построено несколько моделей, основанных на рекуррентном обучении с подкреплением, для решения задачи портфельной оптимизации:

1. Обобщена для портфеля инструментов модель RRL [1]. Обобщенная модель допускает как покупку, так и продажу каждого из активов.

2. В полученную модель добавлено сглаживание оптимизируемого функционала с целью уменьшения дисперсии шумной среды.

3. К итоговой модели применена идея [2], учитывающая непостоянность среды и предполагающая наличие в ней сменяющихся режимов.

Проведено сравнение данных моделей между собой и с общеизвестными моделями портфельной оптимизации (равновзвешенный портфель, портфель с минимальной дисперсией и портфель Марковица) на дневных данных государственных облигаций и глобальных индексов: S&P 500, DAX, Nikkei 225, S&P GSCI и других. Стоит заметить, что при обучении и тестировании моделей учитывалась комиссия, взымаемая за совершение сделок.

Ниже приведена таблица сравнения коэффициентов Шарпа (чем выше значение коэффициента, тем лучше модель) на тестовом периоде (780 торговых дней):

Портфель	Коэффициент Шарпа
RRL модель	**1.136**
сглаженная RRL модель	**1.432**
RRL модель со сменой режимов	**0.827**
равновзвешенный портфель	0.508
портфель с минимальной дисперсией	0.539
портфель Марковица	0.749

Литература

1. Gold C. FX trading via recurrent reinforcement learning. // Proceedings of IEEE International Conference on Computational Intelligence for Financial Engineering, pp. 363–370 (2003), doi:10.1109/CIFER.2003.1196283
2. Maringer D. and Ramtohul T. Regime-Switching Recurrent Reinforcement Learning in Automated Trading. // Natural Computing in Computational Finance, 2011, vol. 4, pp. 100–128.
3. Moody J., Wu L., Liao Y., Saffell M. Performance Functions and Reinforcement Learning for Trading Systems and Portfolios. // Journal of Forecasting, Volume 17, Pages 441–470, 1998