Отчет по конкурсу “Молодая математика России” за 2023 год

Пучкин Никита Андреевич

1 Результаты исследований

Пусть $X, X_1, \dots, X_n$ – независимые одинаково распределенные случайные векторы в $\mathbb{R}^d$ с нулевым математическим ожиданием и ковариационной матрицей Σ. Были исследованы свойства выборочной ковариационной матрицы

$$\widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} X_i X_i^\top .$$

В частности, получен новый, более точный результат, о концентрации выборочной ковариационной матрицы около своего математического ожидания. Предполагалось, что размерность пространства d может быть значительно превосходить объем выборки n, но при этом эффективный ранг Σ, определяемый согласно формуле

$$\mathrm{r}(\Sigma) = \frac{\mathrm{Tr}(\Sigma)}{\|\Sigma\|},$$

оставался небольшим. Целью исследований было получение безразмерных (то есть зависящих только от эффективного ранга Σ, а не от размерности всего пространства d) теоретических верхних оценок на величину $\|\widehat{\Sigma} - \Sigma\|_{\mathrm{F}}^2$. Без ограничения общности считалось, что матрица Σ являлась невырожденной. Тогда свойства $\|\widehat{\Sigma} - \Sigma\|_{\mathrm{F}}^2$ зависят от поведения случайного вектора $\xi = \Sigma^{-1/2} X$. Было сформулировано следующее предположение.

Предположение 1.1 *Существует такая константа $\omega > 0$, что случайный вектор $\xi = \Sigma^{-1/2} X$ удовлетворяет неравенству*

$$\left\| \xi^\top V \xi - \mathrm{Tr}(V) \right\|_{\psi_1} \leqslant \omega \|V\|_{\mathrm{F}} \quad \text{для всех } V \in \mathbb{R}^{d \times d}. \tag{1}$$

Через $\|\cdot\|_{\psi_1}$ обозначена норма Орлича. Напомним, что для произвольной случайной величины η ее норма Орлича $\|\eta\|_{\psi_1}$ определяется как

$$\|\eta\|_{\psi_1} = \inf \left\{ t > 0 : \mathbb{E} e^{|\eta|/t} \leqslant 2 \right\}.$$

Можно привести достаточно примеров распределений, удовлетворяющих Предположению 1.1. В частности, Предположение 1.1 выполнено для случайных векторов $\xi = \Sigma^{-1/2} X$, удовлетворяющих неравенству Хэнсона-Райта. К таким относятся, например, случайные векторы с независимыми субгауссовскими компонентами и случайные векторы, удовлетворяющие логарифмическому неравенству Соболева. Сформулируем главный результат, полученный в ходе исследований.

Теорема 1.2 *Пусть выполнено Предположение 1.1. Зафиксируем $\delta \in (0,1)$ и предположим, что*

$$\mathrm{r}(\Sigma)^2 \left(\mathrm{r}(\Sigma)^2 + \log^2(1/\delta)\right) = o(n), \quad \text{при } n \to \infty.$$

Тогда

$$\left| \left\| \widehat{\Sigma} - \Sigma \right\|_{\mathrm{F}}^2 - \mathbb{E} \left\| \widehat{\Sigma} - \Sigma \right\|_{\mathrm{F}}^2 \right| = O\left(\frac{\|\Sigma\|^2}{n} \max\left\{ \mathrm{r}(\Sigma^2)\sqrt{\log(2/\delta)}, \log(2/\delta) \right\} \right).$$

Таким образом, квадрат нормы Фробениуса $(\widehat{\Sigma} - \Sigma)$ концентрируется около своего математического ожидания. Кроме того, при сформулированном предположении нетрудно показать, что $\mathbb{E} \left\| \widehat{\Sigma} - \Sigma \right\|_{\mathrm{F}}^2$ удовлетворяет следующему неравенству.

Лемма 1.3 *Пусть выполнено Предположение 1.1. Тогда*

$$(\mathrm{Tr}(\Sigma))^2 - \mathrm{Tr}(\Sigma^2) \leqslant n\mathbb{E} \left\| \widehat{\Sigma} - \Sigma \right\|_{\mathrm{F}}^2 \leqslant (\mathrm{Tr}(\Sigma))^2 + (16\omega^2\sqrt{3} - 1)\ \mathrm{Tr}(\Sigma^2).$$

Теорема 1.2 и Лемма 1.3 позволяют установить безразмерные верхнюю и нижнюю оценки для $\left\| \widehat{\Sigma} - \Sigma \right\|_{\mathrm{F}}^2$, которые существенно улучшают результаты, известные из работы [1].

2 Опубликованные и поданные в печать работы

В отчетном году было опубликовано или подано в печать в рецензируемых трудах конференций и научных изданиях пять работ.

- N. Puchkin and V. Ulyanov. Inference via Randomized Test Statistics. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 59(3), pages 1508–1529, 2023.

- D. Belomestny, A. Naumov, N. Puchkin, and S. Samsonov. Simultaneous Approximation of a Smooth Function and its Derivatives by Deep Neural Networks with Piecewise-polynomial Activations. *Neural Networks*, volume 161, pages 242–253, 2023.

- N. Puchkin and V. Shcherbakova. A Contrastive Approach to Online Change Point Detection. *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics (AISTATS), Proceedings of Machine Learning Research*, volume 206, pages 5686–5713, 2023.

- N. Puchkin and N. Zhivotovskiy. Exploring Local Norms in Exp-concave Statistical Learning. *Proceedings of Thirty Sixth Conference on Learning Theory (COLT), Proceedings of Machine Learning Research*, volume 195, pages 1993–2013, 2023.

- N. Puchkin, V. Spokoiny, E. Stepanov, and D. Trevisan. Reconstruction of Manifold Embeddings into Euclidean Spaces via Intrinsic Distances. *ESAIM: Control, Optimization and Calculus of Variations (в печати).*

3 Участие в конференциях и летних школах

В 2023 году выступил с докладом либо представил постер на двух конференциях.

- The 26th International Conference on Artificial Intelligence and Statistics (AISTATS, постер).

- Конференция “Fall into ML 2023”, Москва, 26–28 октября 2023 г. (устный доклад).

Принял участие в качестве слушателя в одной конференции и одной летней школе.

- Летняя школа “Statistics and Learning Theory”, Цахкадзор, Армения, 9–15 июля 2023 г.

- Конференция “Meeting in Mathematical Statistics”, Люмини, Франция, 18–22 декабря 2023 г.

4 Работа в научных центрах

Являюсь научным сотрудником Международной лаборатории стохастических алгоритмов и анализа многомерных данных НИУ ВШЭ и старшим научным сотрудником Института проблем передачи информации им. А. А. Харкевича РАН.

5 Педагогическая деятельность

В 2023 году было принято участие в четырех курсах в качестве лектора или семинариста.

- Онлайн-методы машинного обучения, МФТИ, весенний семестр, лектор.

- Статистическая теория машинного обучения, МФТИ, осенний семестр, лектор.

- Математическая статистика, НИУ ВШЭ, весенний семестр, семинарист.

- High-dimensional probability and statistics, НИУ ВШЭ, весенний семестр, семинарист.

6 Итоги

По итогам исследований за 3 года опубликовано либо подано в печать 5 работ в рецензируемых научных изданиях и 3 статьи в рецензируемых трудах конференций. Результаты были представлены в виде устного доклада или в виде постера на 9 конференциях и летних школах. Также было принято участие в 6 конференциях и летних школах в качестве слушателя. Подготовлена диссертационная работа[1] на тему “Построение оптимальных оценок с помощью метода адаптивных весов в задачах обучения с размеченными и неразмеченными данными”, по результатам защиты которой решением диссертационного совета НИУ ВШЭ по математике 30 мая 2023 г. автору отчета присуждена степень кандидата математических наук.

Проект исследований, представленный при подаче конкурсной заявки, включал в себя:

- разработку и теоретический анализ алгоритма восстановления гладкого многообразия $\mathcal{M}^*$ в $\mathbb{R}^D$, способного адаптироваться к неизвестной размерности $\mathcal{M}^*$;

[1] Текст работы и отзывы членов диссертационного комитета доступны по ссылке https://www.hse.ru/sci/diss/803976128.

- получение новой информационно-теоретической нижней оценки на точность восстановления гладкого многообразия в $\mathbb{R}^D$ в терминах метрики Хаусдорфа при строго ортогональном шуме;

- анализ свойств минимизатора эмпирического риска

$$\widehat{\mathcal{M}} \in \operatorname*{argmin}_{\mathcal{M} \in \mathscr{M}_{\varkappa}^d} \sum_{t=1}^{n} d^2(Y_t, \mathcal{M}),$$

  где $\mathscr{M}_{\varkappa}^d$ – некоторый класс гладких многообразий, в случаях, когда $X_1, \ldots, X_n$ является эргодической и неэргодической марковской цепью.

Решения второй и третьей задач представлены в отчетах за 2022 и 2021 годы соотвественно. В ходе решения первой задачи выяснилось, что адаптация алгоритма к неизвестной размерности восстанавливаемого гладкого многообразия требует оценки на рич этого многообразия. Таким образом, подбор дискретного параметра (размерности многообразия) заменялся на трудоемкую с вычислительной точки зрения задачу оценки непрерывного параметра (рича многообразия). Поэтому было решено отказаться от этого направления исследований ввиду его нецелесообразности. Вместо этого была рассмотрена задача, описанная в первой главе настоящего отчета.

Список литературы

[1] F. Bunea and L. Xiao. On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA. *Bernoulli*, 21(2):1200–1230, 2015.