Компактное представление положения в автокодировщиках нейронной активности

М. А. Оробец*, Н. А. Поспелов‡, В. В. Плюснин§, К. В. Анохин¶

Активность нейронных популяций описывают кодом малой размерности, и от метода зависит, что из этого кода удаётся прочитать. На синтетических записях кальциевой активности мы сравнили автокодировщик с линейными методами и методами обучения многообразий. Когда клетки кодируют только положение животного, лучше всех его передаёт UMAP; когда в активности есть и другие сигналы, при 8 измерениях положение точнее всего читается из кода автокодировщика линейным декодером, а когда клеток, кодирующих сигнал, не связанный с поведением, больше, чем клеток места, положение по его коду из 5 измерений предсказывается точнее, чем по кодам остальных методов.

Ключевые слова: автокодировщик, снижение размерности, клетки места, синтетические данные.

Постановка задачи. Активность популяции нейронов описывают кодом малой размерности: линейными методами (PCA, FA, ICA), методами обучения многообразий (Isomap, UMAP) и автокодировщиками. Мы проверили на синтетических записях, где известно, к чему селективна каждая клетка, как преимущество автокодировщика при малом числе измерений зависит от других сигналов в активности.

Данные. 48 синтетических записей по 60 мин (20 кадров/с) построены генератором driada [1]: 3 популяции по 16 записей, 130–1000 клеток. По-

* *Оробец Маргарита Алексеевна* — лаборант, Федеральный центр мозга и нейротехнологий ФМБА России; автор для переписки, e-mail: margaret.orobets@gmail.com, ORCID: 0009-0002-4231-5329.

Orobets Margarita Alekseevna — Lab Assistant, Federal Center for Brain and Neurotechnologies FMBA of Russia; corresponding author.

‡ *Поспелов Никита Андреевич* — инженер-исследователь, Институт перспективных исследований мозга МГУ имени М. В. Ломоносова (ИПИМ МГУ), e-mail: nik-pos@yandex.ru, ORCID: 0000-0001-6637-2120.

Pospelov Nikita Andreevich — Research Engineer, Institute for Advanced Brain Studies, Lomonosov Moscow State University (IABS MSU).

§ *Плюснин Виктор Владимирович* — лаборант-исследователь, ИПИМ МГУ, e-mail: witkax@mail.ru, ORCID: 0000-0001-9784-8283.

Plusnin Viktor Vladimirovich — Research Assistant, IABS MSU.

¶ *Анохин Константин Владимирович* — директор ИПИМ МГУ, д.м.н., академик РАН, e-mail: k.anokhin@gmail.com, ORCID: 0000-0003-4437-6002.

Anokhin Konstantin Vladimirovich — Director, IABS MSU; DSc (Medicine), Full Member of the Russian Academy of Sciences.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

пуляции: клетки места (все клетки); смешанная (40 % клеток места, 20 % направления головы, по 10 % скорости, скрытой переменной — медленного нестационарного случайного процесса, не связанного с поведением, клеток с несколькими переменными и молчащих); с преобладанием скрытой переменной (30 % клеток места, 40 % скрытой переменной, 30 % молчащих). Кальциевый сигнал — спайки, свёрнутые с экспонентой (постоянная затухания 2 с), плюс шум.

Методы. Каждый метод строил код размерности 5, 8 и 12 по первым 80 % записи без их последних 30 с; оценочной частью служили последние 20 %. Активность нейронов приводилась к $[0, 1]$; нормировка только по обучающей части меняла среднюю корреляцию не более чем на 0.02. Параметры подбирались на отдельной когорте записей: у Isomap и UMAP 100 и 50 соседей из 15–200, у автокодировщика скрытый слой 64 нейрона, Adam (10^{-3}), до 400 эпох. Мерой качества служила корреляция предсказания переменной по коду (линейной регрессией или методом k ближайших соседей с равными весами, k из 5, 15, 50, 150 выбиралось на проверочной части обучающих данных) с истинной переменной на оценочной части; для положения корреляции с координатами x и y усреднялись. Сравнения парные по 16 записям (критерий Уилкоксона) с поправкой Бонферрони на все 90 сравнений ($p < 0.00056$).

Результаты. Когда клетки кодировали только положение, лучше всех его линейно передавал UMAP (при 5 измерениях 0.87 против 0.69 у автокодировщика; рис. 1, а), а с нелинейным декодером все методы при 5–12 измерениях давали корреляцию 0.93–0.95. В популяциях с другими сигналами UMAP терял положение (при 8 измерениях 0.58 и 0.23; рис. 1, б, в). В смешанной популяции при 8 измерениях автокодировщик передавал положение линейным декодером точнее всех методов без нейросетей (0.72 против 0.66 у PCA и ICA, во всех 16 записях, $p < 0.001$); стандартное отклонение его корреляции между запусками с разной инициализацией составляло в среднем 0.03, меньше этого перевеса. С нелинейным декодером FA достигал того же уровня (0.83 у FA против 0.85 у автокодировщика): положение есть в обоих кодах, но из кода автокодировщика его извлекает уже линейный декодер. Isomap в этой популяции уступал PCA (0.57 против 0.66) из-за переноса вложения на новые кадры: если вложение строилось по всей записи, Isomap давал 0.63. В популяции с преобладанием скрытой переменной при 5 измерениях автокодировщик опережал все методы и с линейным декодером (0.58 против 0.25 у PCA и ICA), и с нелинейным (0.65 против 0.37 у UMAP), во всех 16 записях ($p < 0.001$). При 12 измерениях перевес автокодировщика над FA и PCA становился незначимым.

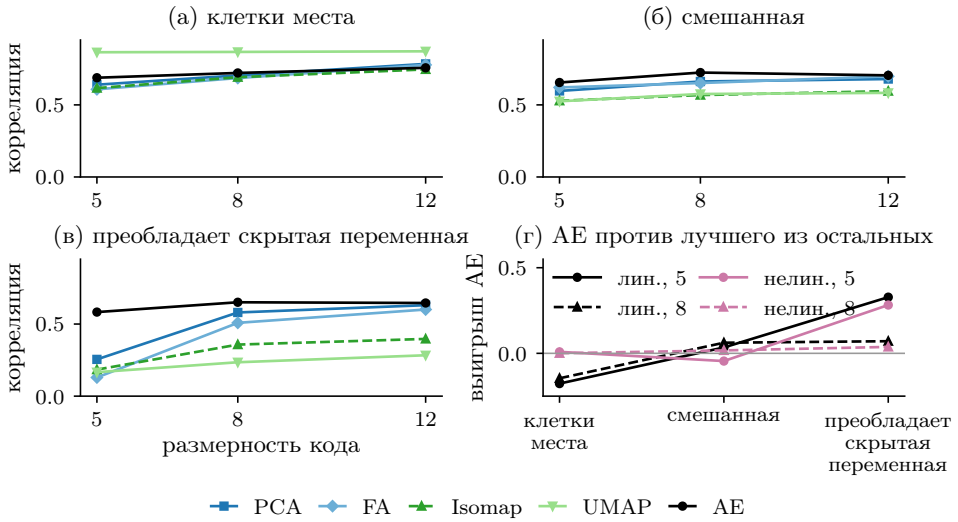


Рис. 1: Корреляция предсказания положения на оценочной части, среднее по 16 записям. (а)–(в) Линейный декодер: (а) клетки места; (б) смешанная популяция; (в) преобладание скрытой переменной. (г) Выигрыш AE (автокодировщика) над лучшим из PCA, FA, Isomap и UMAP при 5 и 8 измерениях, линейный и нелинейный декодер. ICA совпадает с PCA.

Обсуждение. Активность нейронных популяций лежит на многообразиях малой размерности [2, 3]. Выбор метода зависит от того, что ещё есть в активности: если популяция кодирует одну переменную, её многообразие лучше всего восстанавливают методы обучения многообразий. Если в активности есть и другие сигналы, в том числе не связанные с поведением, автокодировщик сохраняет положение в коде малой размерности: линейные методы и UMAP при 8 измерениях передают его линейному декодеру хуже, а при преобладании скрытой переменной и 5 измерениях теряют. Если так же устроены и реальные записи, где закодированные переменные известны не полностью, то отсутствие переменной в коде линейного метода малой размерности ещё не означает её отсутствия в активности популяции.

Список литературы

[1] N. Pospelov, V. Plusnin, O. Rogozhnikova, A. Ivanova, V. Sotskov, M. Orobets, K. Toropova, O. Ivashkina, V. Avetisov, K. Anokhin, “DRIADA: an open-source Python toolkit for linking single-neuron selectivity and population activity to behavior”, 2026, arXiv: 2607.00851. DOI: 10.48550/arXiv.2607.00851.

[2] J. A. Gallego, M. G. Perich, L. E. Miller, S. A. Solla, "Neural Manifolds for the Control of Movement", *Neuron*, **94**:5 (2017), 978–984. DOI: 10.1016/j.neuron.2017.05.025.

[3] E. Altan, S. A. Solla, L. E. Miller, E. J. Perreault, "Estimating the dimensionality of the manifold underlying multi-electrode neural recordings", *PLOS Computational Biology*, **17**:11 (2021), e1008591. DOI: 10.1371/journal.pcbi.1008591.

Статья поступила 9 ноября 2026 г.

Compact representation of position in autoencoders of neural activity

M. A. Orobets, N. A. Pospelov, V. V. Plusnin, K. V. Anokhin

Neural population activity is described by low-dimensional codes, and the method decides what can be read from the code. On synthetic calcium imaging recordings we compared an autoencoder with linear and manifold learning methods. When cells encode only the animal's position, UMAP conveys it best; when the activity carries other signals too, position is read most accurately from the autoencoder's code by a linear decoder at 8 dimensions, and when cells encoding a signal unrelated to behavior outnumber place cells, position is predicted more accurately from its 5-dimensional code than from the codes of the other methods.

Keywords: autoencoder, dimensionality reduction, place cells, synthetic data.

References

[1] N. Pospelov, V. Plusnin, O. Rogozhnikova, A. Ivanova, V. Sotskov, M. Orobets, K. Toropova, O. Ivashkina, V. Avetisov, K. Anokhin, "DRIADA: an open-source Python toolkit for linking single-neuron selectivity and population activity to behavior", 2026, arXiv: 2607.00851. DOI: 10.48550/arXiv.2607.00851.

[2] J. A. Gallego, M. G. Perich, L. E. Miller, S. A. Solla, "Neural Manifolds for the Control of Movement", *Neuron*, **94**:5 (2017), 978–984. DOI: 10.1016/j.neuron.2017.05.025.

[3] E. Altan, S. A. Solla, L. E. Miller, E. J. Perreault, "Estimating the dimensionality of the manifold underlying multi-electrode neural recordings", *PLOS Computational Biology*, **17**:11 (2021), e1008591. DOI: 10.1371/journal.pcbi.1008591.

Received on November 9, 2026