

Оценка внутренней размерности SVD-представлений языков

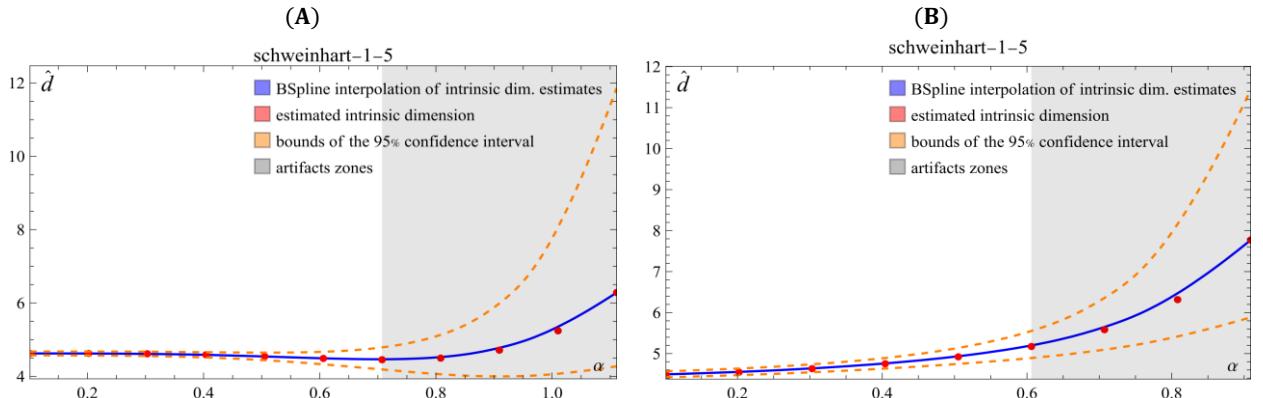


Рисунок J1. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 5$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 5$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

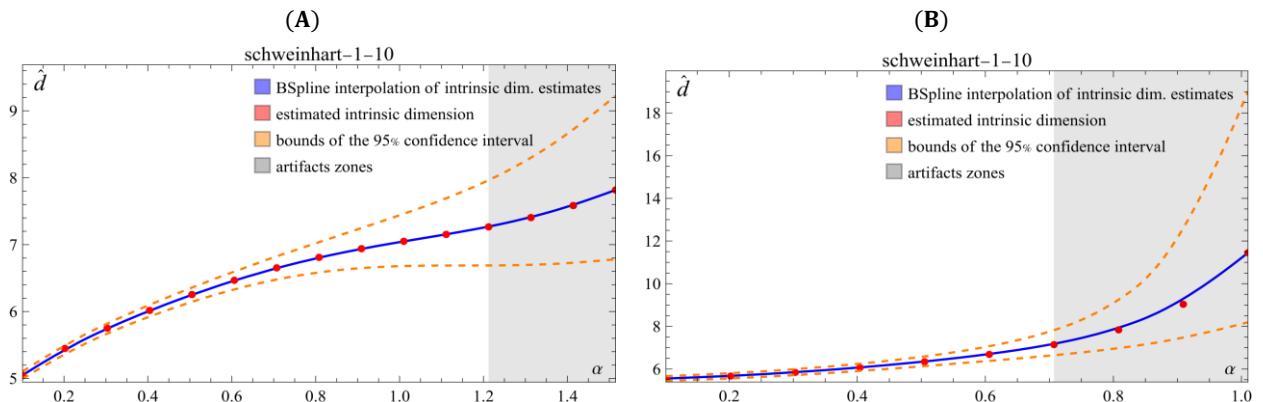


Рисунок J2. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 10$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 10$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

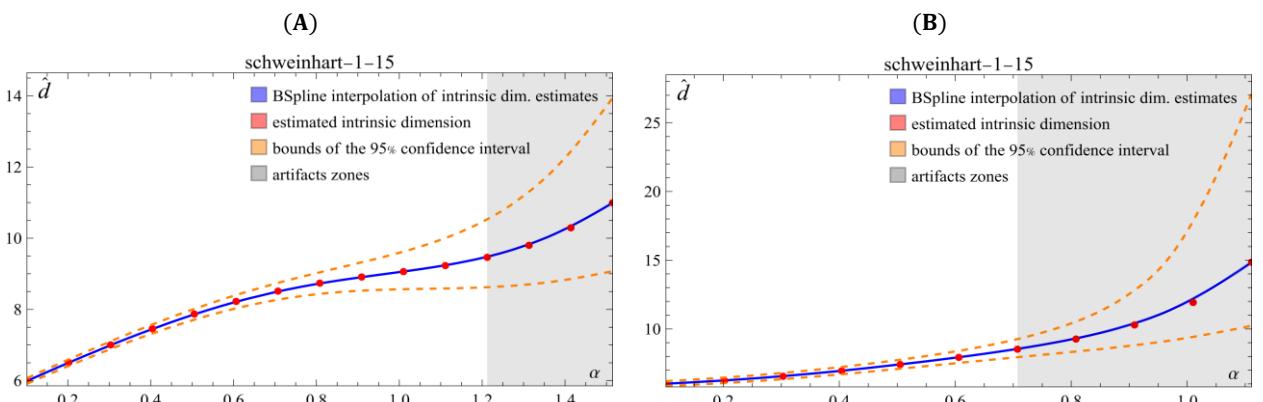


Рисунок J3. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 15$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 15$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

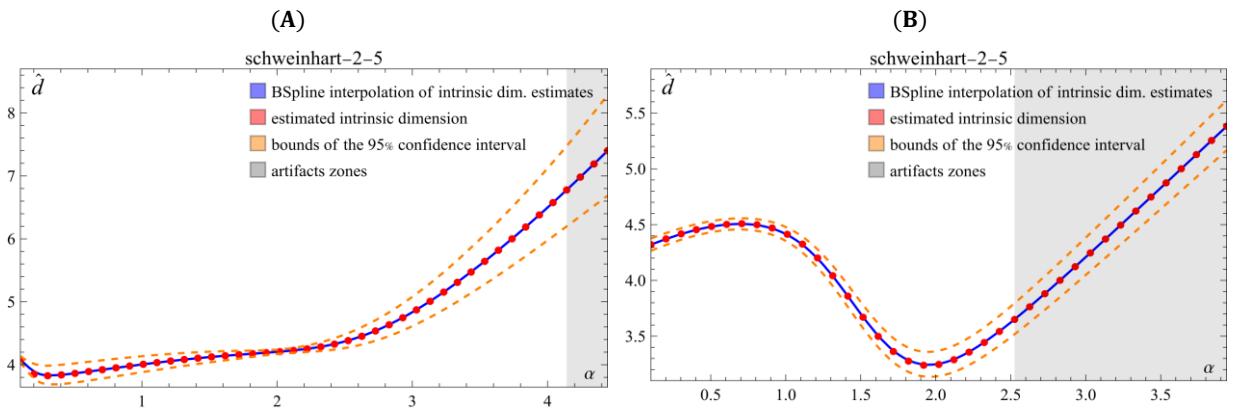


Рисунок J4. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 5$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 5$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

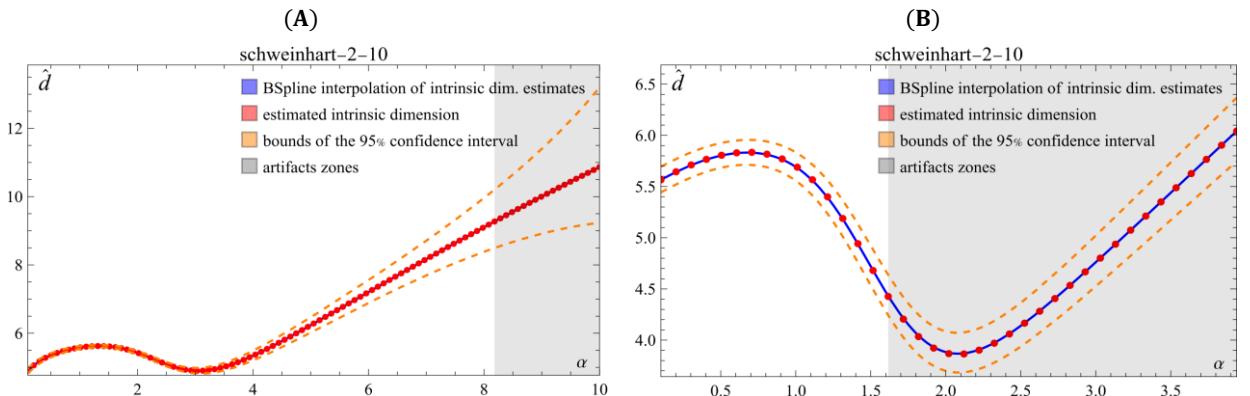


Рисунок J5. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 10$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 10$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

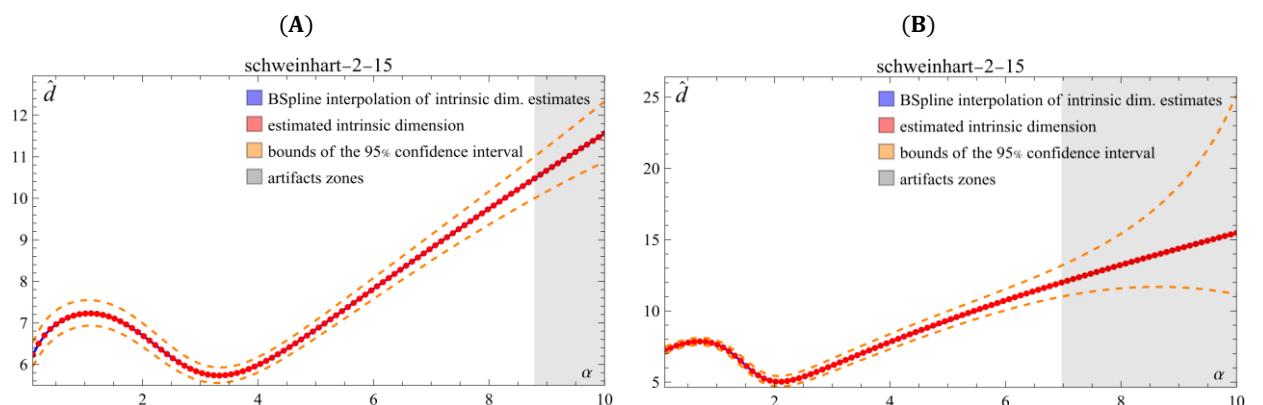


Рисунок J6. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 15$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 15$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

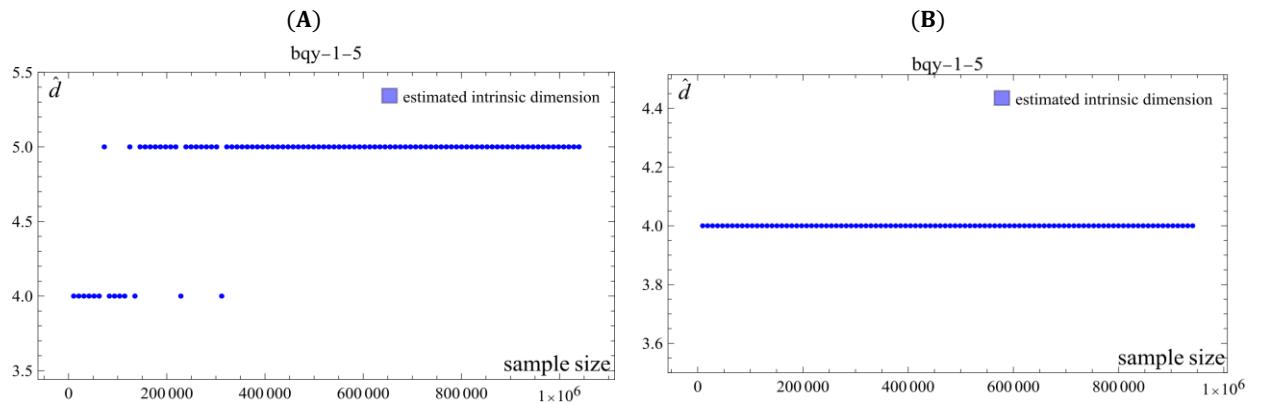


Рисунок J7. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 5$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 5$).

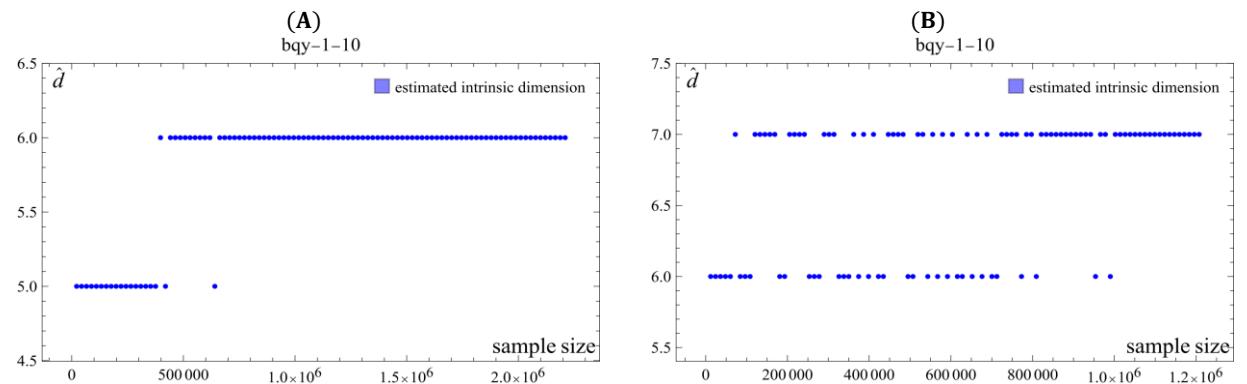


Рисунок J8. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 10$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 10$).

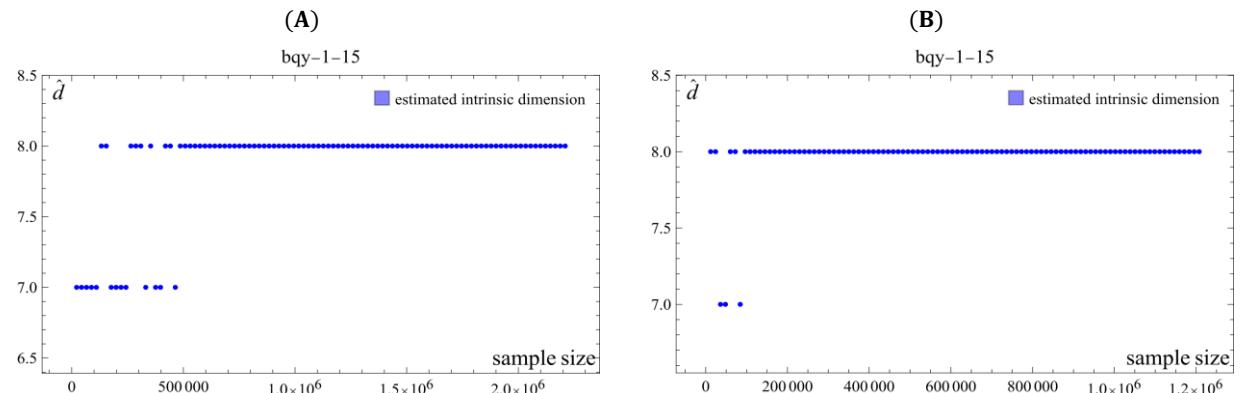


Рисунок J9. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 1, d = 15$); (B) SVD-векторные представления английской национальной литературы ($n = 1, d = 15$).

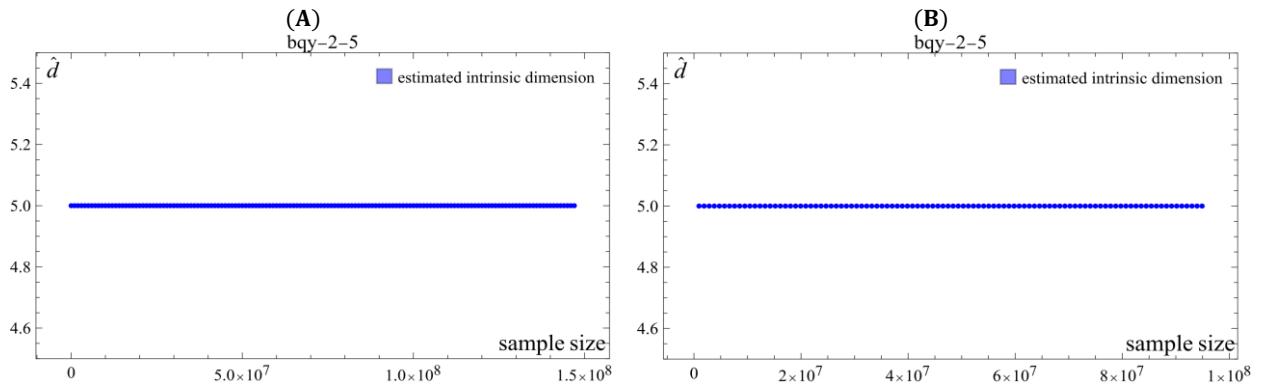


Рисунок J10. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 5$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 5$).

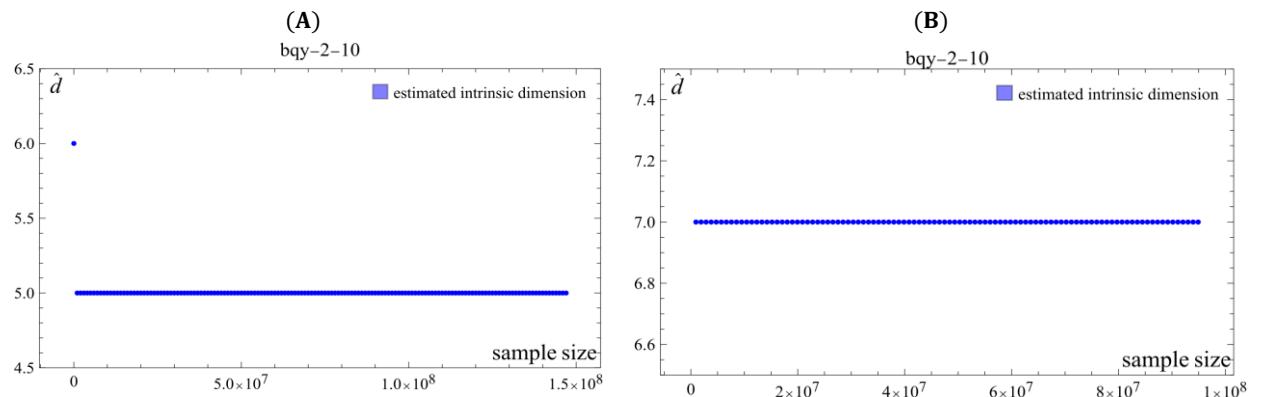


Рисунок J11. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 10$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 10$).

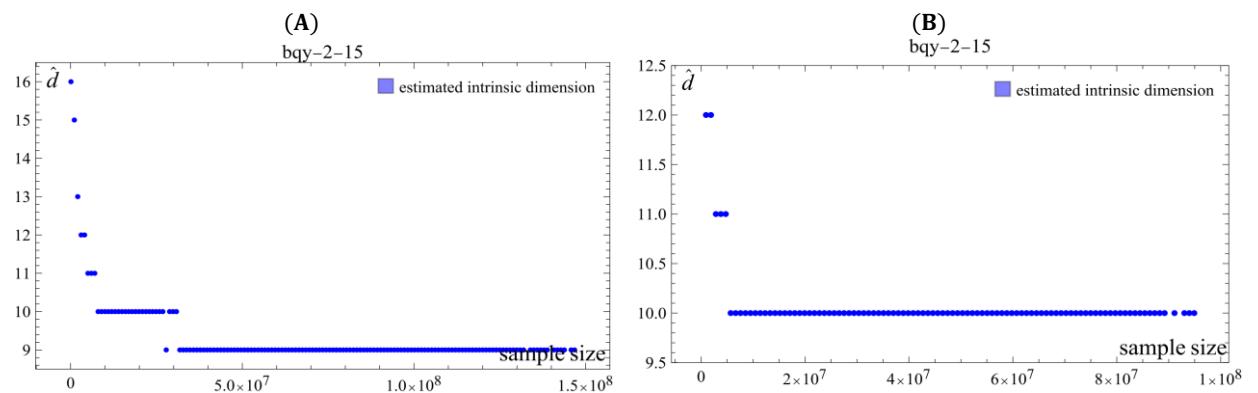


Рисунок J12. Зависимость оценок алгоритма Brito от параметра α для: (A) SVD-векторные представления русской национальной литературы ($n = 2, d = 15$); (B) SVD-векторные представления английской национальной литературы ($n = 2, d = 15$).

Оценка внутренней размерности CBOW-представлений языков

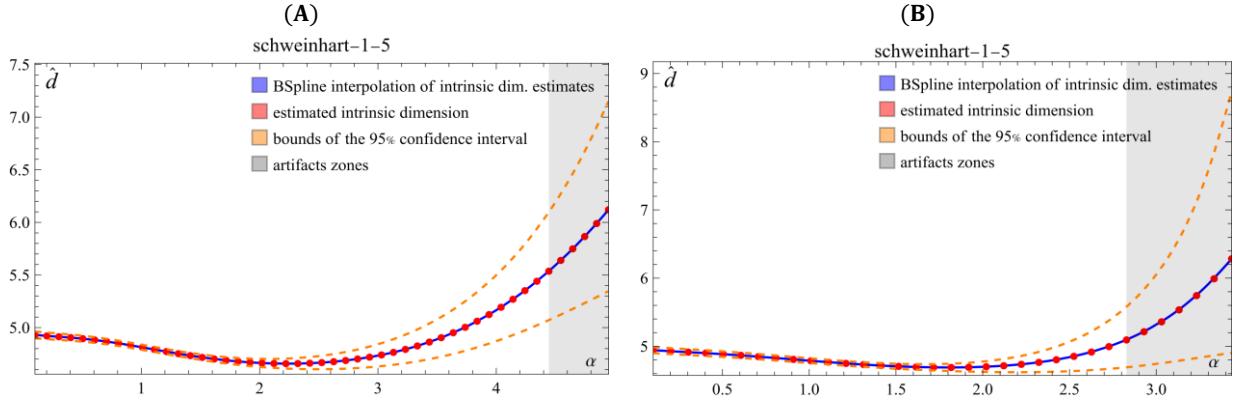


Рисунок К1. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 1, d = 5$); (B) CBOW – векторные представления английской национальной литературы ($n = 1, d = 5$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

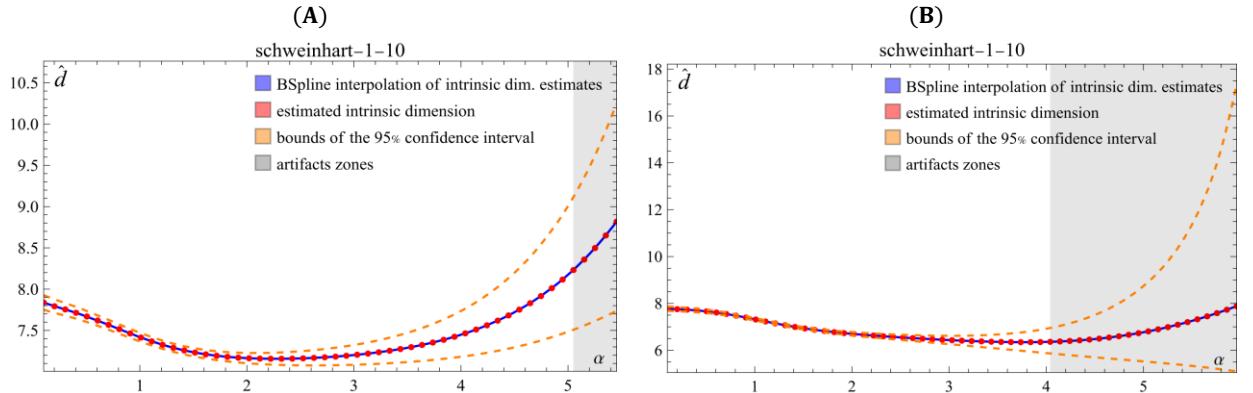


Рисунок К2. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 1, d = 10$); (B) CBOW – векторные представления английской национальной литературы ($n = 1, d = 10$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

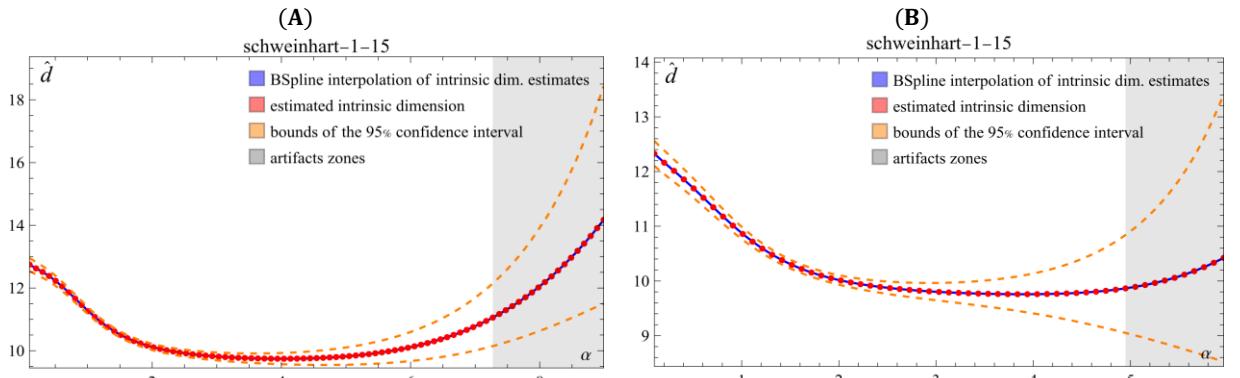


Рисунок К3. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 1, d = 15$); (B) CBOW – векторные представления английской национальной литературы ($n = 1, d = 15$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

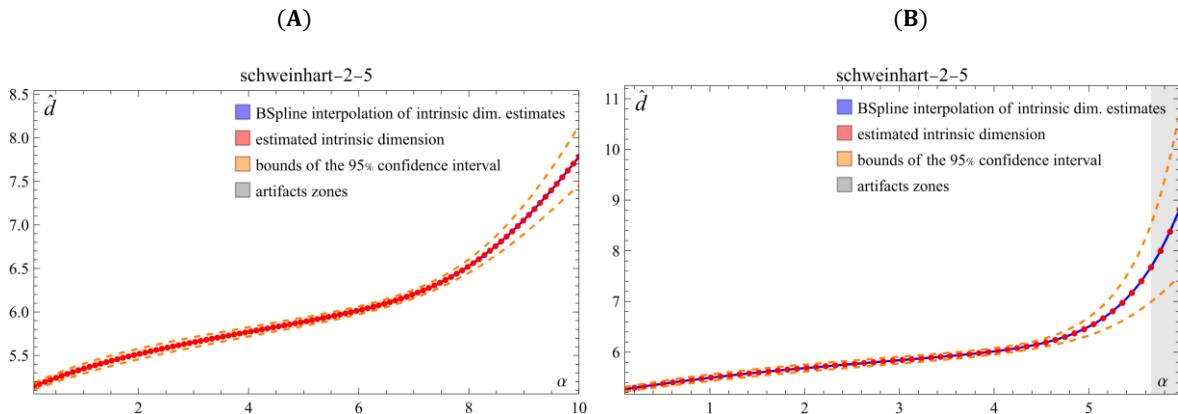


Рисунок К4. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 2, d = 5$); (B) CBOW – векторные представления английской национальной литературы ($n = 2, d = 5$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

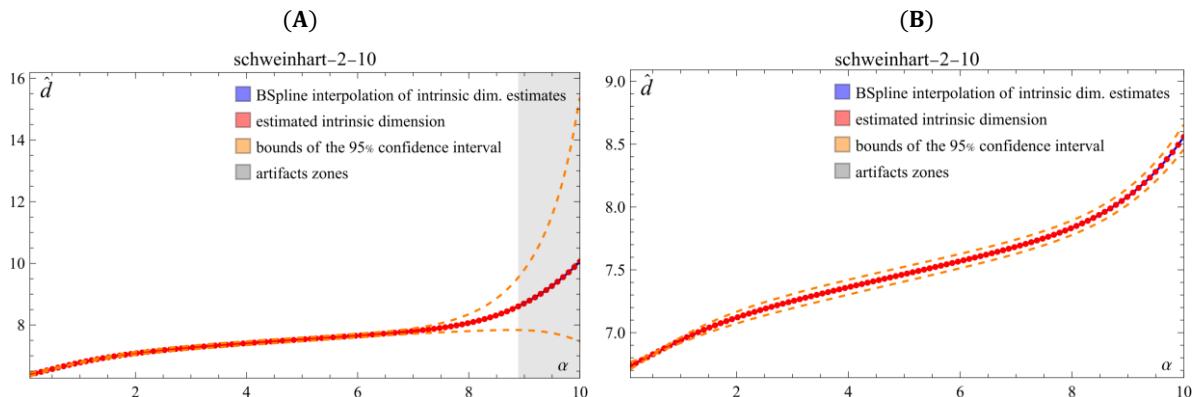


Рисунок К5. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 2, d = 10$); (B) CBOW – векторные представления английской национальной литературы ($n = 2, d = 10$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

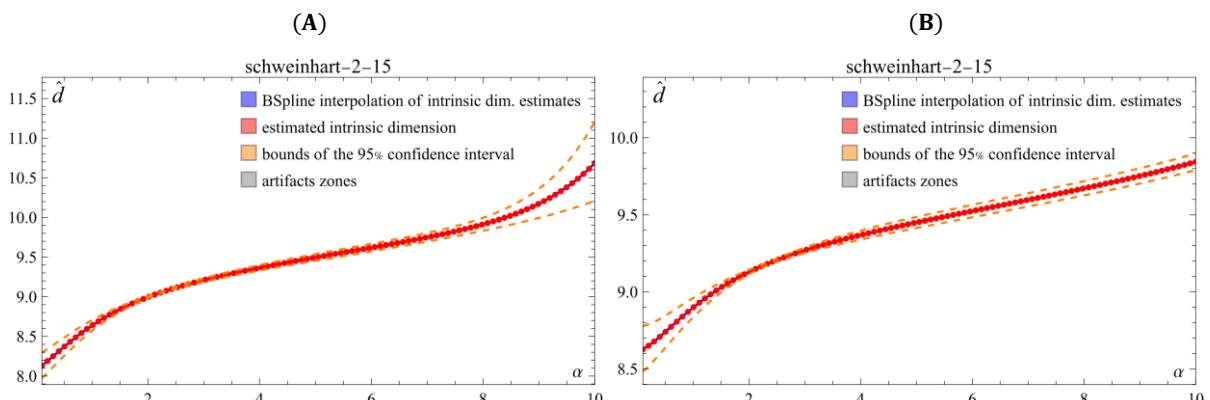


Рисунок К6. Зависимость оценок алгоритма Schweinhart от параметра α для: (A) CBOW – векторные представления русской национальной литературы ($n = 2, d = 15$); (B) CBOW – векторные представления английской национальной литературы ($n = 2, d = 15$). Красные точки представляют собой оценки алгоритма, синяя линия — интерполяция B-spline зависимости оценок от α , оранжевая пунктирная линия обозначает верхние и нижние границы 95%-го доверительного интервала. На рисунке также визуализированы критерии информативности параметра α , такие как расхождение доверительных интервалов (серыми зонами).

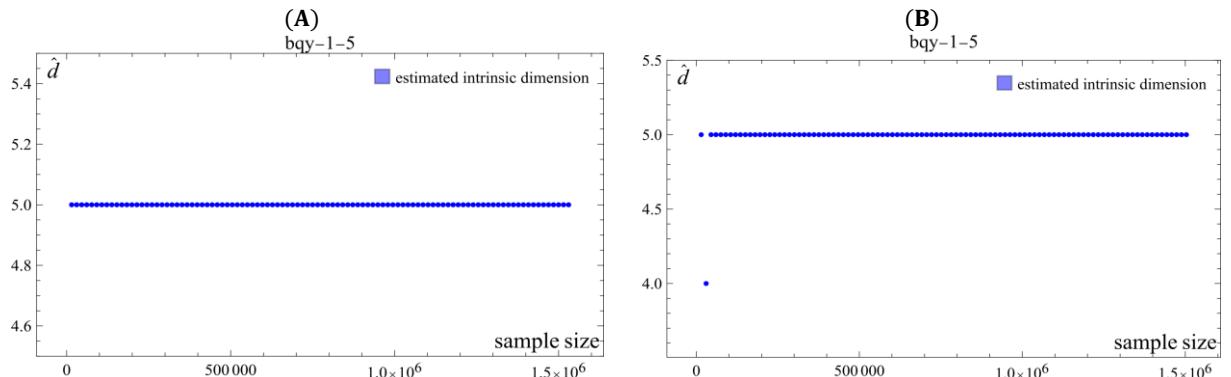


Рисунок K7. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 1, d = 5$); (B) CBOW-векторные представления английской национальной литературы ($n = 1, d = 5$).

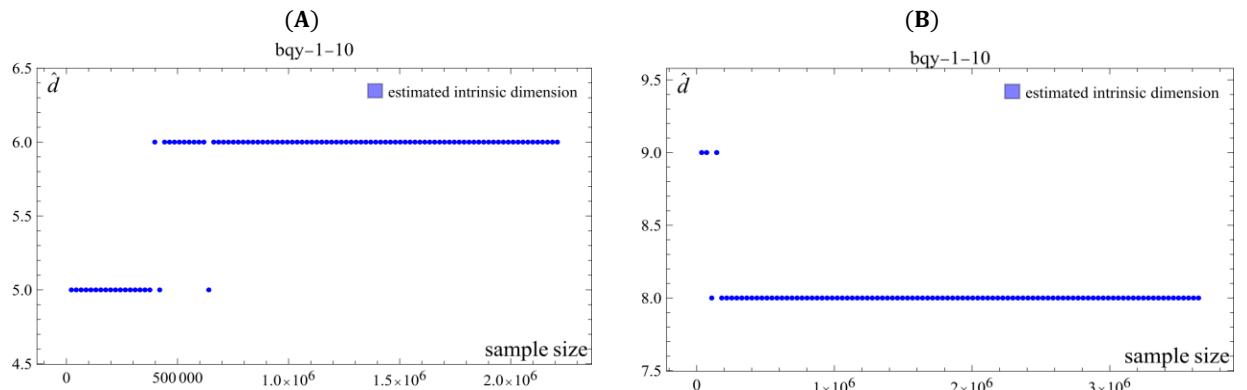


Рисунок K8. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 1, d = 10$); (B) CBOW-векторные представления английской национальной литературы ($n = 1, d = 10$).

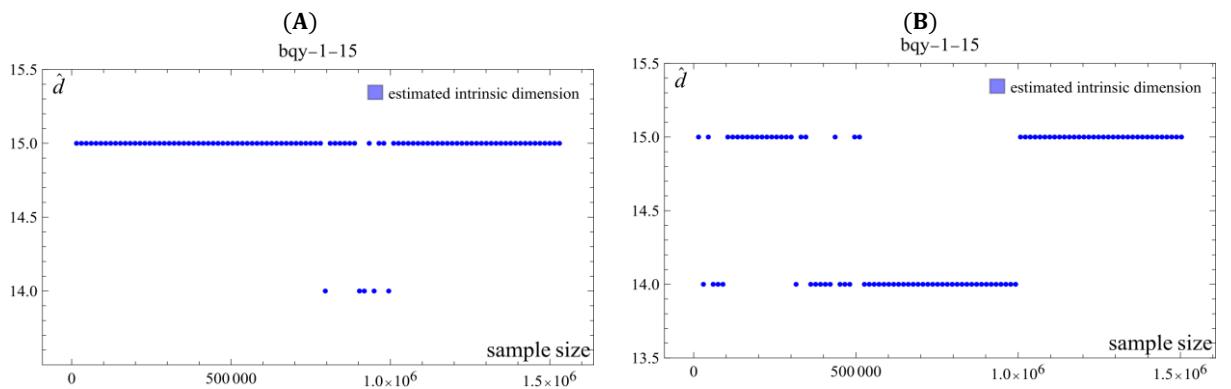


Рисунок K9. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 1, d = 15$); (B) CBOW-векторные представления английской национальной литературы ($n = 1, d = 15$).

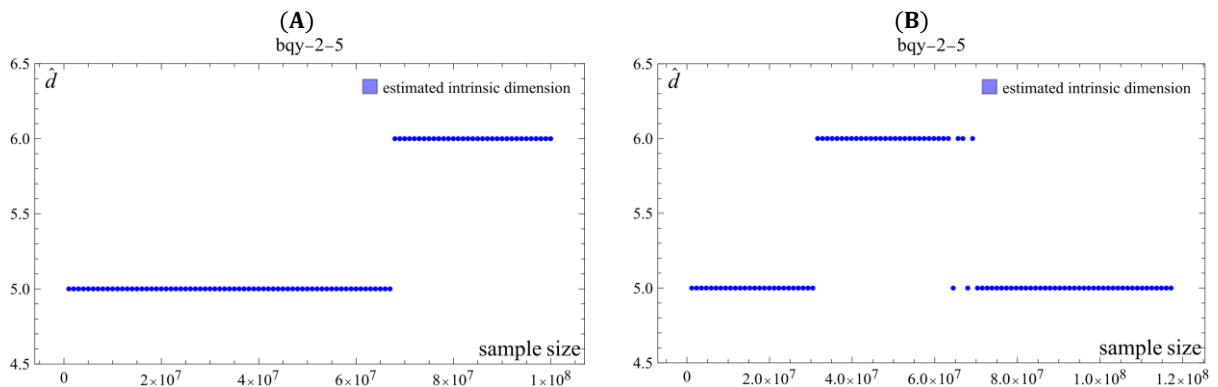


Рисунок K10. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 2, d = 5$); (B) CBOW-векторные представления английской национальной литературы ($n = 2, d = 5$).

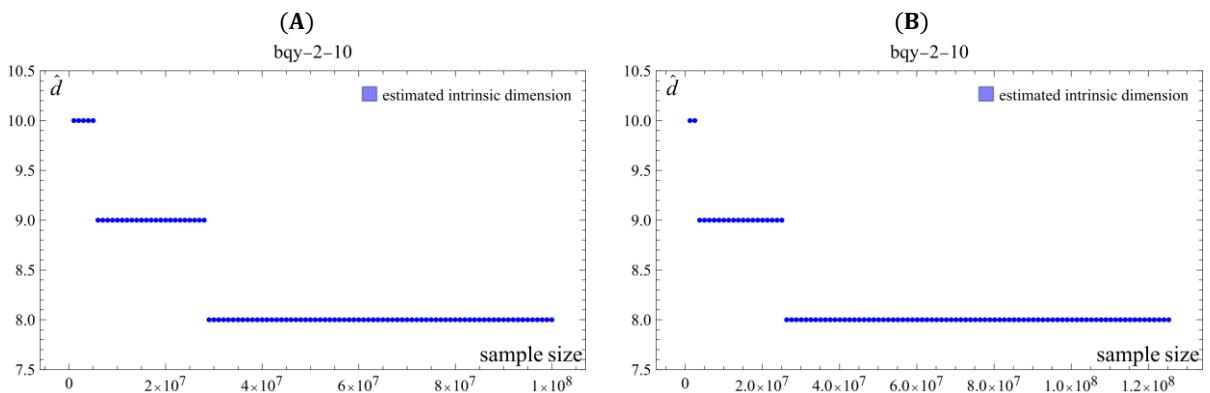


Рисунок K11. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 2, d = 10$); (B) CBOW-векторные представления английской национальной литературы ($n = 2, d = 10$).

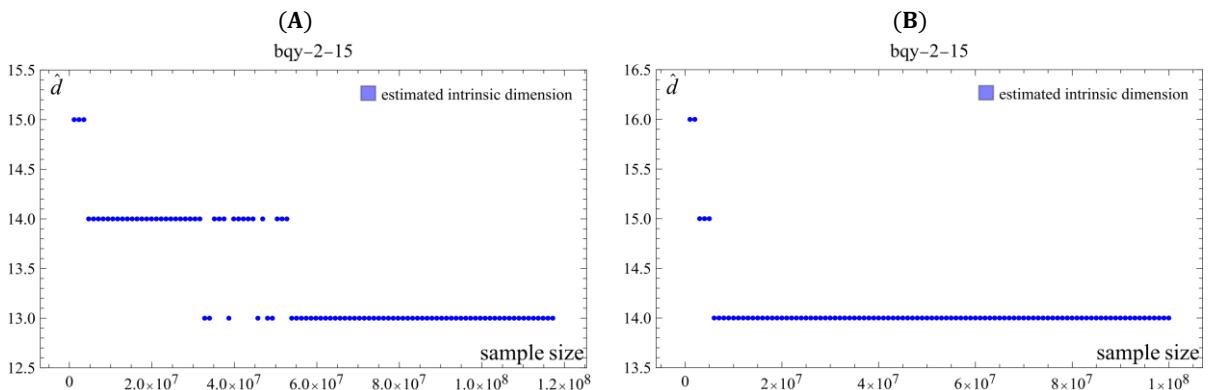


Рисунок K12. Зависимость оценок алгоритма Brito от параметра α для: (A) CBOW-векторные представления русской национальной литературы ($n = 2, d = 15$); (B) CBOW-векторные представления английской национальной литературы ($n = 2, d = 15$).