

## Вопросы к коллоквиуму по Многомерному анализу данных, осень-2015

1. Многомерное нормальное распределение. Вектор мат.ож. и ковар.матрица при лин. преобразовании (умножении на матрицу).
2. Оценки вектора средних и ковар.матрицы. Несмещенная оценка ковар. матрицы
3. Распределение вектора средних.
4. Переход к новым признакам с помощью ортогональной матрицы. Пример про способности по математике и физике (выписать матрицу вращения).
5. Разложение матрицы данных при переходе к новым признакам в виде суммы и в матричном виде.
6. Как определяется вклад новых признаков.
7. Сингулярное разложение, как строится.
8. Сингулярное разложение. В каком смысле оно единственно.
9. Разложение Шмидта.
10. Выборочный анализ главных компонент и сингулярное разложение, общее и различия.
11. Анализ главных компонент на генеральном языке как частный случай разложения Шмидта
12. Почему главные компоненты так называются, в каком смысле они главные.
13. Оптимальность сингулярного разложения в смысле аппроксимации матрицей ранга  $r$
14. Оптимальность сингулярного разложения в смысле аппроксимации подпространством размерности  $r$
15. Оптимальность в анализе главных компонент в статистической терминологии (через дисперсии).
16. Оптимизация в АГК в терминах ковариационных матриц.
17. В двух статистических пакетах получились разные главные компоненты. Отчего так могло получиться?
18. Смысл первой ГК, если все ковариации (корреляции) исходных признаков положительны.
19. Разница между АГК по корреляционной и по ковариационной матрице на примере двух признаков. Когда что использовать.
20. Способы выбора числа главных компонент.
21. Почему доля собственного числа по отношению к сумме собственных чисел называется объясненной долей общей дисперсии?
22. На основе каких элементов сингулярного разложения интерпретируются главные компоненты как линейные комбинации исходных признаков? Привести формулу и пример.
23. АГК с точки зрения построения базиса в пространстве индивидов и в пространстве признаков. Координаты в новых базисах.
24. Как выявить индивидов, которые плохо описываются плоскостью первых двух главных компонент?
25. Как вычислить значения главных компонент для индивида, которого не было в исходной выборке. А как вычислить значения факторных значений?
26. В каком случае координаты в ортонормированном базисе можно назвать корреляциями?
27. Чему равны суммы по строкам и по столбцам в матрице, составленной из собственных векторов в АГК?
28. Чему равны суммы по строкам и по столбцам в матрице факторных нагрузок в АГК?
29. Как интерпретировать скалярное произведение строк в матрице факторных нагрузок в АГК?
30. Как нарисовать исходные орты в плоскости первых двух главных компонент?
31. Зачем и когда первые две координаты факторных нагрузок рисуются в единичном круге?
32. Чему равна норма  $i$ -го вектора из главных компонент?
33. Как формализовать веса для признаков и для индивидов в АГК?
34. Какова модель в факторном анализе?
35. Что делает АГК в модели факторного анализа при равных общностях?
36. Какая разница между АГК и факторным анализом?
37. Связь между числом факторов и числом признаков для корректности задачи.
38. Что минимизируется в методе MINRES? В чем разница с тем, что минимизируется в АГК?
39. Какой вид имеет функция правдоподобия в ФА?
40. Проверка значимости модели ФА.
41. Критерий сферичности Бартлетта, для чего нужен.
42. Что такое общность и уникальность признака? Какие факторы не находит факторный анализ?
43. Общность как множественный коэффициент корреляции.
44. Как интерпретируются признаки в ФА?
45. Зачем нужны вращения в ФА? Как устроены ортогональные вращения?
46. Вращение по методу varimax.
47. Методы нахождения факторных значений, LS и WLS (метод Бартлетта).
48. Факторная структура (корреляции исходных признаков с факторами) и факторный паттерн (коэффициенты лин. комбинации, с которыми исходные признаки выражаются через факторы) в случае ортогональных и неортогональных факторов.