**Метод канонических корреляций**

В таблице (файл “Данные” лист “Показатели”) приведены значения показателей производственной деятельности для 53 предприятий машиностроения. Использовались следующие показатели: Y1 - производительность труда; Y2 – индекс снижения себестоимости продукции; Y3 – рентабельность; X1 – трудоемкость единицы продукции; X2 – удельный вес рабочих в составе ППП; X3 – удельный вес покупных изделий; X4 – коэффициент сменности оборудования; X5 –премии и вознаграждения на одного работника; X6 – удельный вес потерь от брака; X7 – фондоотдача; X8 – среднегодовая численность ППП; X9 – среднегодовая стоимость ОПФ; X10 – среднегодовой фонд заработной платы ППП; X11 – фондовооруженность труда; X12 – оборачиваемость нормируемых оборотных средств; X13 - оборачиваемость ненормируемых оборотных средств; X14 – непроизводственные расходы.

Требуется для выбранных в соответствии с вариантом факторов (смотри таблицу 1) провести канонический анализ для двух групп факторов Y и X:

1. Найти оценки канонических переменных (найти канонические веса, использующиеся для вычисления значений канонических переменных) и канонических корреляций.
2. Произвести оценку значимости полученных канонических корреляций и, соответственно, отсеять незначимые пары канонических переменных. Записать выражения для значимых канонических переменных через исходные признаки.
3. Найти корреляции между каноническими переменными и переменными из каждого множества Y и X (координаты векторов канонических нагрузок).
4. Вычислить извлеченную дисперсию каждой канонической переменной и совокупностью канонических переменных (для каждого множества) и определить избыточность каждого множества исходных данных.

**Задание 1**

Найдем оценки канонических переменных (найдем канонические веса, использующиеся для вычисления значений канонических переменных) и канонических корреляций.

Для начала импортируем исходные данные в Wolfram Mathematica. Получим две группы величин Y и X. Стандартизируем величины каждой из групп для дальнейшей работы.

Найдем выборочные матрицы ковариации A11, A22, A12, A21. Они имеют следующий вид:

Задача для нахождения одной из канонических переменных была сведена к задаче на нахождение собственных значений и собственных векторов следующей системы:

у которой были найдены собственные значения λ2 и соответствующие им собственные векторы . Далее выбираем наибольшие собственное значение λ2и собственный вектор, соответствующий этому значению. Нормируем выбранный вектор, согласно условию:

После находим вектор из условия:

Таким образом, были найдены все канонические веса, для построения первой пары канонических переменных:

Значение для канонической корреляции для этой пары будет равняться .

Аналогично находим оставшиеся пары канонических переменных.

**Задание 2**

Произведем оценку значимости полученных канонических корреляций и, соответственно, отсеем незначимые пары канонических переменных. Запишем выражения для значимых канонических переменных через исходные признаки.

Для проверки статистической значимости полученных канонических корреляций воспользуемся критерием отношения правдоподобия. Гипотезы будут выглядеть следующим образом:

Наблюдаемое значение статистики имеет распределение , где .

Записав отношение правдоподобия для q=1,2,3, было получено, что при q=3 p-уровень значимости равняется 0,99. Таким образом, количество значимых пар канонических переменных равняется двум.

Запишем значимые канонические переменные через исходные признаки:

**Задание 3**

Найдем корреляции между каноническими переменными и переменными из каждого множества Y и X (координаты векторов канонических нагрузок).

Для расчета координат векторов канонических нагрузок воспользуемся формулами:

**Задание 4**

Вычислим извлеченную дисперсию каждой канонической переменной и совокупностью канонических переменных (для каждого множества) и определим избыточность каждого множества исходных данных.

Вычислить данные показатели можно последующим формулам.

Доля абсолютной дисперсии признаков объясняемая переменной равняется ||.

Доля абсолютной дисперсии признаков объясняемая переменной равняется ||.

Относительная доля дисперсии признаков объясняемая переменной равняется

Относительная доля дисперсии признаков объясняемая переменной равняется

Относительная доля дисперсии признаков объясняемая совокупностью переменных равняется .

Относительная доля дисперсии признаков объясняемая совокупностью переменных равняется .

Избыточность первого множества относительно второго .

Избыточность второго множества относительно первого .

Таблица 1. Абсолютная доля дисперсии признаков, объясняемая каноническими переменными

|  |  |  |  |
| --- | --- | --- | --- |
| Каноническая переменная | Абсолютная доля дисперсии признаков | Каноническая переменная | Абсолютная доля дисперсии признаков |
| u1 | 1.268 | v1 | 1.521 |
| u2 | 0.983 | v2 | 0.810 |

Таблица 2. Относительная доля дисперсии признаков, объясняемая каноническими переменными

|  |  |  |  |
| --- | --- | --- | --- |
| Каноническая переменная | Относительная доля дисперсии признаков | Каноническая переменная | Относительная доля дисперсии признаков |
| u1 | 0,423 | v1 | 0,254 |
| u2 | 0.328 | v2 | 0.135 |
| Сумма | 0,75 |  | 0,38 |

Избыточность первого множества относительно второго равняется 0,63.

Избыточность второго множества относительно первого равняется 0,33.