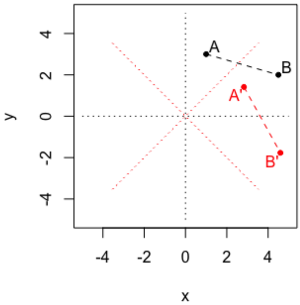
Матрицы и определители. Факторный анализ

Н.П.Алексеева

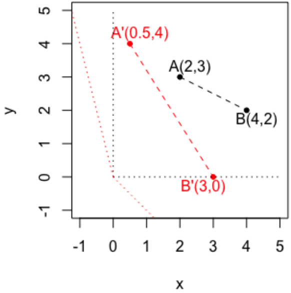
СПбГУ, мат-мех ф-т

Матрицы и векторы



$$\overbrace{\begin{cases} ax+by = x' \\ cx+dy = y' \end{cases}}^{\text{система уравнений}} \iff \overbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}^{\text{матрица}} \cdot \overbrace{\begin{bmatrix} x \\ y \end{bmatrix}}^{\text{вектор}} = \overbrace{\begin{bmatrix} x' \\ y' \end{bmatrix}}^{\text{вектор}}$$

Умножение матрицы на вектор



$$\begin{cases} x - 0.5y = x' \\ -x + 2y = y' \end{cases} \iff \begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$A(2,3) \to A'(\ ,\ )?\quad B(4,2) \to B'(\ ,\ )?$

Умножение матрицы на вектор

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 - 0.5 \cdot 3 \\ -2 + 2 \cdot 3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 - 0.5 \cdot 2 \\ -4 + 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Умножение матрицы на матрицу

$$\begin{cases} ax_1 + by_1 = x'_1 \\ cx_1 + dy_1 = y'_1 \end{cases} \quad \begin{cases} ax_2 + by_2 = x'_2 \\ cx_2 + dy_2 = y'_2 \end{cases} \quad \ldots \begin{cases} ax_k + by_k = x'_k \\ cx_k + dy_k = y'_k \end{cases}$$

Если нужно получить преобразования сразу нескольких точек с координатами (x_1, y_1),..., (x_k, y_k), то эти координаты можно собрать в матрицу

$$\begin{bmatrix} x_1 & \ldots & x_k \\ y_1 & \ldots & y_k \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x_1 & \ldots & x_k \\ y_1 & \ldots & y_k \end{bmatrix} = \begin{bmatrix} x'_1 & \ldots & x'_k \\ y'_1 & \ldots & y'_k \end{bmatrix}$$

Умножение матрицы на матрицу. Пример

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 - 0.5 \cdot 3 \\ -2 + 2 \cdot 3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 - 0.5 \cdot 2 \\ -4 + 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 0.5 & 3 \\ 4 & 0 \end{bmatrix}$$

Умножение матрицы размерности m, n на n-мерный вектор

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = y_1 \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = y_m \end{cases}$$

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix},$$

$$A[m,n]X[n,1] = Y[m,1] \quad \text{или} \quad AX = Y$$

Умножение матрицы размерности m, n на матрицу размерности n, k

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & \dots & x_{1k} \\ \vdots & & \\ x_{n1} & \dots & x_{nk} \end{bmatrix},$$

$$A[m, n]X[n, k] = Y[m, k] \quad \text{или} \quad AX = Y$$

$$Y = \begin{bmatrix} y_{11} & \dots & y_{1k} \\ \vdots & & \\ y_{m1} & \dots & y_{mk} \end{bmatrix}.$$

Транспонирование

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \dots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad X^T = [x_1, \dots, x_n]$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad X^T = [x_1, \ldots, x_n]$$

$$XX^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot [x_1, \ldots, x_n] = \begin{bmatrix} x_1x_1 & \ldots & x_1x_n \\ \vdots & \ldots & \vdots \\ x_nx_1 & \ldots & x_nx_n \end{bmatrix}$$

Транспонирование произведения матриц

$(AB)^T = B^T A^T$

$$\begin{bmatrix} 1 & -0.5 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 0.5 & 3 \\ 4 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -0.5 & 2 \end{bmatrix} = \begin{bmatrix} 0.5 & 4 \\ 3 & 0 \end{bmatrix}$$

Единичные матрицы

Матрица $\mathbb{I}$ с единицами на главной диагонали и с нулями вне ее называется единичной.

Для любой матрицы A произведение $A\mathbb{I} = \mathbb{I}A = A$.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

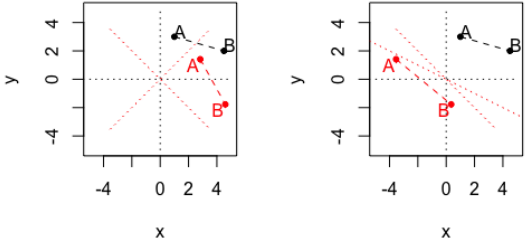
Обратные и ортогональные матрицы

Если $\mathrm{AB} = \mathbb{I}$, то $\mathrm{B} = \mathrm{A}^{-1}$ называется обратной.

Если $\mathrm{A}^{-1} = \mathrm{A}^{\mathrm{T}}$, то матрица A называется ортогональной ($\mathrm{AA}^{\mathrm{T}} = \mathrm{A}^{\mathrm{T}}\mathrm{A} = \mathbb{I}$).

При ортогональном преобразовании $\mathrm{Y} = \mathrm{AX}$,
$\mathrm{X} = \mathrm{A}^{-1}\mathrm{Y} = \mathrm{A}^{\mathrm{T}}\mathrm{Y}$, расстояния между точками сохраняются.

$$\sum_{\mathrm{i}} \mathrm{x}_{\mathrm{i}}^2 = \mathrm{X}^{\mathrm{T}}\mathrm{X} = (\mathrm{A}^{\mathrm{T}}\mathrm{Y})^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{Y} = \mathrm{Y}^{\mathrm{T}}\mathrm{A}\mathrm{A}^{\mathrm{T}}\mathrm{Y} = \mathrm{Y}^{\mathrm{T}}\mathrm{Y} = \sum_{\mathrm{j}} \mathrm{y}_{\mathrm{j}}^2 .$$



$$C_1 = \begin{bmatrix} 0.71 & 0.71 \\ -0.71 & 0.71 \end{bmatrix}, \qquad C_2 = \begin{bmatrix} 0.71 & -1.41 \\ -0.71 & 0.71 \end{bmatrix}$$

$$C_1^T C_1 = \mathbb{I}, \qquad C_2^T C_2 \neq \mathbb{I}\,.$$

Собственные числа и собственные вектора

Для каждой квадратной матрицы Σ можно найти вектор x такой, что $\Sigma \mathrm{x} = \lambda \mathrm{x}$, где λ некоторая константа.

$$\Sigma = \begin{bmatrix} 1 & \varrho \\ \varrho & 1 \end{bmatrix}, \quad \mathrm{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \Sigma \mathrm{x} = \begin{bmatrix} 1+\varrho \\ 1+\varrho \end{bmatrix} = (1+\varrho) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Поскольку свойство $\Sigma \mathrm{x} = \lambda \mathrm{x}$ справедливо для любых векторов вида $\mathrm{x} = \begin{bmatrix} \mathrm{c} \\ \mathrm{c} \end{bmatrix}$, то выбираем из них представителя с единичной нормой, то есть $\mathrm{c}^2 + \mathrm{c}^2 = 1$, $\mathrm{c} = \frac{1}{\sqrt{2}}$.
Параметр λ называется собственным числом, а вектор x собственным вектором.

Иллюстрация собственных чисел и собственных векторов

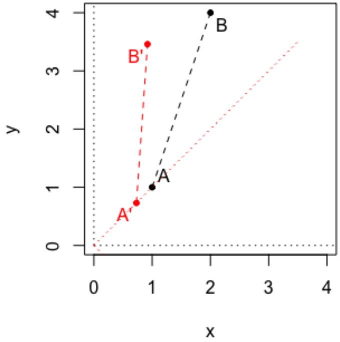
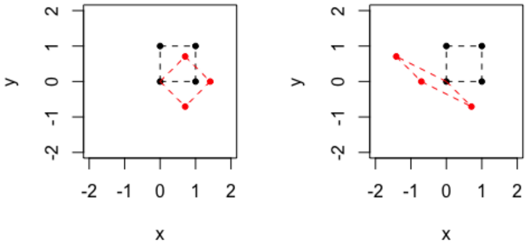


Figure: Преобразование при помощи матрицы $\begin{bmatrix} 1 & \varrho \\ \varrho & 1 \end{bmatrix}$, $\varrho = -0.27$.

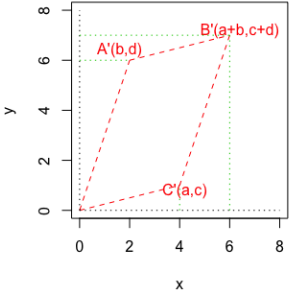
Определитель как площадь



$$C_1 = \begin{bmatrix} 0.71 & 0.71 \\ -0.71 & 0.71 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0.71 & -1.41 \\ -0.71 & 0.71 \end{bmatrix}$$

$$\det(C_1) = 1, \qquad \det(C_2) = -0.5, .$$

Вычисление определителя



$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = (a{+}b)(c{+}d) - \frac{ac}{2} - \frac{bd}{2} - \frac{c(2b+a)}{2} - \frac{b(2c+d)}{2} = ad - bc$$

Применение определителя в решении систем уравнений

$$\begin{cases} ax + by = 0 \\ cx + dy = 0 \end{cases} \iff \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$$

$$y = -\frac{a}{b}x = -\frac{c}{d}x \iff ad = bc$$

$$\begin{cases} a_{11}x_1 + \ldots + a_{1k}x_k = 0 \\ \vdots \\ a_{k1}x_1 + \ldots + a_{kk}x_k = 0 \end{cases} \iff \begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & \ldots & \vdots \\ a_{k1} & \ldots & a_{kk} \end{vmatrix} = 0$$

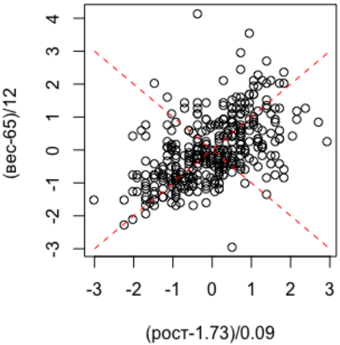
Вычисление собственных чисел

$$\Sigma \mathrm{x} = \lambda \mathrm{x} \iff (\Sigma - \lambda \mathbb{I})\mathrm{x} = \mathbb{O} \iff |\Sigma - \lambda \mathbb{I}| = 0$$

$$\Sigma = \begin{bmatrix} 1 & \varrho \\ \varrho & 1 \end{bmatrix}, \quad \mathrm{x} = \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \end{bmatrix}$$

$$\begin{vmatrix} 1-\lambda & \varrho \\ \varrho & 1-\lambda \end{vmatrix} = 0 \iff (1-\lambda)^2 - \varrho^2 = 0, \quad \lambda = \begin{cases} 1+\varrho \\ 1-\varrho \end{cases}$$

При $\lambda = 1 + \varrho$ получаем собственный вектор $\mathrm{A}_1 = \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ из уравнения $-\varrho \mathrm{x}_1 + \varrho \mathrm{x}_2 = 0$, второй собственный вектор $\mathrm{A}_2 = \begin{bmatrix} \mathrm{x}_1 \\ -\mathrm{x}_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ при $\lambda = 1 - \varrho$ получаем из уравнения $\varrho \mathrm{x}_1 + \varrho \mathrm{x}_2 = 0$.



$$Y_1 = \frac{1}{\sqrt{2}}\left(\frac{\text{рост}-1.73}{0.09} + \frac{\text{вес}-65}{12}\right), \quad Y_2 = \frac{1}{\sqrt{2}}\left(-\frac{\text{рост}-1.73}{0.09} + \frac{\text{вес}-65}{12}\right)$$

- L продолжительность жизни
- M количество чиновников
- P процент бедных
- A число автомобилей
- V объемы продаваемой водки

	L	M	P	A	V
1	68.90	1060.00	7.80	5.50	25.30
2	68.10	1101.00	9.50	15.30	28.00
3	67.60	1147.00	10.10	30.20	30.00
4	69.20	1204.00	10.00	44.50	23.50
5	69.20	1602.00	9.80	58.60	18.00
6	64.60	1893.00	5.50	93.30	38.40
7	67.00	2777.00	6.20	122.00	29.60

Выявление ведущих факторов

Первой главной компонентой Y_1 признаков $X_1, \ldots, X_k$ называется линейная комбинация исходных признаков

$$Y_1 = \alpha_{11} X_1 + \ldots + \alpha_{k1} X_k,$$

где коэффициенты $\alpha_1 = (\alpha_{11}, \ldots, \alpha_{k1})^T$ выбраны таким образом, чтобы дисперсия $D(Y_1) = \lambda_1$ была максимальной.
Вторая главная компонента

$$Y_2 = \alpha_{12} X_1 + \ldots + \alpha_{k2} X_k,$$

где коэффициенты $\alpha_2 = (\alpha_{12}, \ldots, \alpha_{k2})^T$ выбраны таким образом, что компоненты Y_1 и Y_2 некоррелированы, а дисперсия $D(Y_2) = \lambda_2$ является максимальной из всех линейных комбинаций, некоррелированных с Y_1. Остальные главные компоненты строятся аналогично.

Вычисление коэффициентов для главных компонент

- Пусть имеются признаки $X_1, \ldots, X_k$, собранные в виде вектора $X = (X_1, \ldots, X_k)^T$ и Σ ковариационная матрица с элементами в виде коэффициентов ковариации $\sigma_{ij} = \mathbb{E}(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)$.
- $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ упорядоченные собственные числа матрицы Σ
- $A_1, \ldots, A_k$ соответствующие собственные вектора, $\Sigma A_i = \lambda_i A_i$.
- $A = [A_1|A_2|\ldots|A_k]$ матрица, составленная из собственных векторов
- Вектор $Y = A^T X$ называется вектором главных компонент.

Функции в R: дисперсии главных компонент

Дисперсия главной компоненты $\mathbb{D}Y_j = \lambda_j$.
Суммарная дисперсия $\sum_{j=1}^{k} \lambda_i = \sum_{i=1}^{k} \mathbb{D}X_i$.

data – таблица со столбцами L,M,P,A,V

$PC < -princomp(scale(data))$
PC[[1]]

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
λ_i	1.76	0.97	0.46	0.20	0.13
$\lambda_i^<$	1.76	2.72	3.18	3.38	3.51
$\frac{\lambda_i}{\sum_j \lambda_j} \cdot 100\%$	50.09	27.56	13.06	5.69	3.60
$\frac{\lambda_i^<}{\sum_j \lambda_j} \cdot 100\%$	50.09	77.65	90.71	96.40	100.00

Факторные нагрузки

Корреляция $\beta_{ij} = \mathrm{cor}(X_i, Y_j) = \frac{\alpha_{ij}\sqrt{\lambda_j}}{\sigma_{ii}}$ между признаком X_i и главной компонентой Y_j называется факторной нагрузкой.

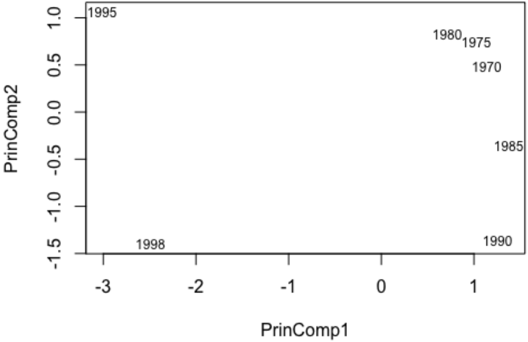
```
PC <-princomp(scale(data))
PC$loadings
```

	Comp.1	Comp.2
L	0.47	-0.38
M	-0.43	-0.54
P	0.48	-0.04
A	-0.45	-0.47
V	-0.41	0.59

- По первой компоненте выделяются годы, когда много бедных с большей продолжительностью жизни, мало чиновников и автомобилей (до и после 1991 года)
- По второй компоненте много водки, мало чиновников и автомобилей (1980,1995 по сравнению с 1990 и 1998).

Двумерная диаграмма факторов

```
plot(PC$scores[, 1], PC$scores[, 2], type = "n", xlab =
"PrinComp1", ylab = "PrinComp2")
text(PC$scores[, 1], PC$scores[, 2], row.names(data), cex = 0.75)
```



Формула восстановления

Коэффициенты $A = \{\alpha_{ij}\}$ выражаются через собственные векторы ковариационной матрицы, $Y = A^T X$, $X = AY$.

$$X_i = \sum_{j=1}^{k} \alpha_{ij} Y_j$$

Figure: Восстановление средней продолжительности жизни по первым двум факторам. $\widehat{L} = (0.896 f_1 + (-0.398 f_2)) \cdot 1.64 + 67.8$.