

Тема 2. Ковариация и собственный вектор

1. Подготовка данных

Временной ряд рассматривается либо как некоторая выборка из генеральной совокупности либо как описание некоторой детерминированной функции.

Все значения каждого признака в различные моменты времени образуют временной ряд, который обозначается вектором

$$x_i = \{x_{1i}^0, x_{2i}^0, \dots, x_{mi}^0\}^T$$

Показатели центра распределения

Среднеарифметические значения временных рядов:

$$\bar{x}_i = \frac{1}{m} \sum_k x_{ki}$$

Мода — это наиболее часто наблюдаемая величина изучаемого временного ряда

Медиана — это значение наблюдения, которое находится в середине ранжированного ряда данных, т.е. наблюдение, занимающее срединное положение

Показатели вариации

Дисперсия – отклонение наблюдаемого значения (для каждого наблюдения) от среднего арифметического (несмещенная оценка) :

$$\sigma_{x_1} = \frac{\sum_{k=1}^m (x_{k1}^0 - \bar{x}_1)(x_{k1}^0 - \bar{x}_1)}{m}$$

Если размер выборки относительно ограничен, то для более точного расчета применяется формула смещенной (исправленной) дисперсии:

$$\sigma_{x_1} = \frac{\sum_{k=1}^m (x_{k1}^0 - \bar{x}_1)(x_{k1}^0 - \bar{x}_1)}{m-1}$$

Показатели вариации

Среднеквадратическое (стандартное) отклонение показывает абсолютное отклонение измеренных значений от среднего арифметического и определяется по формуле :

$$\varepsilon_{x_1} = \sqrt{\frac{\sum_{k=1}^m (x_{k1}^0 - \bar{x}_1)(x_{k1}^0 - \bar{x}_1)}{m}} = \sqrt{\sigma_{x_1}}$$

Показатели вариации

Ковариация двух случайных величин определяется следующим образом (несмещенная оценка)

$$\text{cov}(x_1^0, x_2^0) = \sigma_{x_1 x_2} = \frac{\sum_{k=1}^m (x_{k1}^0 - \bar{x}_1)(x_{k2}^0 - \bar{x}_2)}{m}$$

Если размер выборки относительно ограничен, то для более точного расчета применяется формула смещенной (исправленной) ковариации:

$$\text{cov}(x_1^0, x_2^0) = \sigma_{x_1 x_2} = \frac{\sum_{k=1}^m (x_{k1}^0 - \bar{x}_1)(x_{k2}^0 - \bar{x}_2)}{m - 1}$$

Показатели вариации

Коэффициент корреляции

$$r_{x_1 x_2} = \frac{\sigma_{x_1 x_2}}{\varepsilon_{x_1}^0 \varepsilon_{x_2}^0}$$

Коэффициент корреляции – это безразмерная величина, которая может принимать значения из отрезка :

- ✓ 1 – имеет место абсолютная положительная корреляция между рассматриваемыми величинами
- ✓ -1 – имеет место абсолютная отрицательная корреляция между рассматриваемыми величинами
- ✓ 0 – линейная корреляционная связь отсутствует.

Ковариационная матрица

Пространство признаков описывается матрицей

$$\mathbf{X}^0 = \begin{bmatrix} x_{11}^0 & x_{12}^0 & \boxtimes & x_{1i}^0 & \boxtimes & x_{1n}^0 \\ x_{21}^0 & x_{22}^0 & \boxtimes & x_{2i}^0 & \boxtimes & x_{2n}^0 \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{i1}^0 & x_{i2}^0 & \boxtimes & x_{ii}^0 & \boxtimes & x_{in}^0 \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{m1}^0 & x_{m2}^0 & \boxtimes & x_{mi}^0 & \boxtimes & x_{mn}^0 \end{bmatrix}$$

ИЛИ

$$\mathbf{X}^0 = [x_1^0 \quad x_2^0 \quad \boxtimes \quad x_n^0]$$

Операции с данными

При вычислении ковариационной и корреляционной матриц используются операции вычитания средних (**центрирование**) и деления на стандарты (**нормирование**).

Операция **нормирования** в пространстве соответствует изменению масштабов по всем осям координат так, чтобы величина рассеяния, характеризуемая величиной дисперсии, стала равной единице.

Ковариационная матрица

При вычислении ковариационной матрицы используется операция «центрирование».

$$\mathbf{A} = \frac{1}{m} \mathbf{X}^T \mathbf{X}$$

$$\mathbf{X} = [x_1 \quad x_2 \quad \boxtimes \quad x_n]$$

$$x_{ji} = x_{ji}^0 - \bar{x}_i$$

Среднемесячные курсы изменения валют в 2007 году

Дата	Доллар США, x_1	Евро, x_2	$x_1 - \bar{x}_1$	$x_2 - \bar{x}_2$	$(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)$
Январь	26.53	34.39	0.95	-0.62	-0.58800
Февраль	26.34	34.41	0.76	-0.60	-0.45845
Март	26.11	34.57	0.52	-0.44	-0.22887
Апрель	25.84	34.89	0.26	-0.12	-0.030282
Май	25.82	34.91	0.24	-0.10	-0.024225
Июнь	25.91	34.78	0.33	-0.23	-0.076827
Июль	25.54	35.03	-0.04	0.02	-0.000816
Август	25.62	34.90	0.04	-0.11	-0.004653
Сентябрь	25.33	35.16	-0.25	0.15	-0.036817
Октябрь	24.90	35.39	-0.69	0.38	-0.262545
Ноябрь	24.47	35.91	-1.12	0.90	-1.006731
Декабрь	24.57	35.78	-1.01	0.77	-0.778048
	\bar{x}_1	\bar{x}_2		Сумма	-3.496277
	25.58	35.01		Ковариация	-0.317843

Ковариационная матрица

	Доллар США	Евро
Доллар США	0.4345	-0.318
Евро	-0.318	0.2372

Вычисление корреляционной матрицы

При вычислении корреляционной матрицы используются операции центрирование и нормирование

$$r_{x_1^0 x_2^0} = \frac{\sigma_{x_1 x_2}}{\varepsilon_{x_1^0} \varepsilon_{x_2^0}}$$

$$\mathbf{A}_k = \mathbf{E} \mathbf{A} \mathbf{E}$$

E - диагональная матрица (нормирование)

	Доллар США	Евро
Доллар США	1,0	-0,989
Евро	-0,989	1,0

Матрица вторых моментов

Вычисление вторых моментов

$$\mu_{x_1 x_2} = \frac{\sum_{k=1}^m x_{k1}^0 x_{k2}^0}{m-1}$$

$$\mathbf{A}^0 = \frac{1}{m-1} (\mathbf{X}^0)^T \mathbf{X}^0$$

	Доллар США	Евро
Доллар США	714,35	976,72
Евро	976,72	1337,4

Матрица нормированных вторых моментов

При вычислении матрицы нормированных начальных вторых моментов используется операция нормирование

$$\bar{\mu}_{x_1^0 x_2^0} = \frac{\mu_{x_1^0 x_2^0}}{\varepsilon_{x_1^0} \varepsilon_{x_2^0}}$$

$$\mathbf{A}_k^0 = \mathbf{P} \mathbf{A}^0 \mathbf{P}$$

	Доллар США	Евро
Доллар США	1,0	0,999
Евро	0,999	1,0



Матрица

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \boxtimes & x_{1i} & \boxtimes & x_{1n} \\ x_{21} & x_{22} & \boxtimes & x_{2i} & \boxtimes & x_{2n} \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{i1} & x_{i2} & \boxtimes & x_{ii} & \boxtimes & x_{in} \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{m1} & x_{m2} & \boxtimes & x_{mi} & \boxtimes & x_{mn} \end{bmatrix}$$



Транспонированная матрица

$$\mathbf{X}^{0T} = \begin{bmatrix} x_{11} & x_{21} & \boxtimes & x_{i1} & \boxtimes & x_{m1} \\ x_{12} & x_{22} & \boxtimes & x_{i2} & \boxtimes & x_{m2} \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{1i} & x_{2i} & \boxtimes & x_{ii} & \boxtimes & x_{mi} \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ x_{1n} & x_{2n} & \boxtimes & x_{in} & \boxtimes & x_{mn} \end{bmatrix}$$

Умножение матрицы на вектор

$$\mathbf{c} = \mathbf{Xb}$$

$$c_i = \sum_{k=1}^m x_{ik} b_k$$

$$\begin{Bmatrix} c_1 \\ c_2 \\ \vdots \\ c_i \\ \vdots \\ c_n \end{Bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1i} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2i} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ii} & \cdots & x_{im} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{ni} & \cdots & x_{nm} \end{bmatrix} \begin{Bmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{Bmatrix}$$

СОБСТВЕННЫЙ ВЕКТОР

Пример вектора и собственного вектора

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \times \begin{Bmatrix} 1 \\ 3 \end{Bmatrix} = \begin{Bmatrix} 2 \times 1 + 3 \times 3 \\ 2 \times 1 + 1 \times 3 \end{Bmatrix} = \begin{Bmatrix} 11 \\ 5 \end{Bmatrix}$$

$$\begin{bmatrix} 4 & 3 \\ 3 & 1,5 \end{bmatrix} \times \begin{Bmatrix} 3 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 4 \times 3 + 3 \times 2 \\ 3 \times 3 + 1,5 \times 2 \end{Bmatrix} = \begin{Bmatrix} 18 \\ 12 \end{Bmatrix} = 6 \times \begin{Bmatrix} 3 \\ 2 \end{Bmatrix}$$

$$\begin{bmatrix} 4 & 3 \\ 3 & 1,5 \end{bmatrix} \times \begin{Bmatrix} 2 \\ -3 \end{Bmatrix} = \begin{Bmatrix} 4 \times 2 - 3 \times 3 \\ 3 \times 2 - 1,5 \times 3 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = -0,5 \times \begin{Bmatrix} 2 \\ -3 \end{Bmatrix}$$

СОБСТВЕННЫЙ ВЕКТОР

Уравнение собственных векторов

$$X\mathbf{b} - \lambda\mathbf{b} = 0$$

$$\begin{bmatrix} 4 & 3 \\ 3 & 1,5 \end{bmatrix} \times \begin{Bmatrix} 3 \\ 2 \end{Bmatrix} - 6 \times \begin{Bmatrix} 3 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 18 \\ 12 \end{Bmatrix} - 6 \times \begin{Bmatrix} 3 \\ 2 \end{Bmatrix} = 0$$

Алгоритм вычисления собственных векторов

1. Выбрать произвольное начальное (нулевое) приближение собственного вектора b_0

Положить $k=0$

2. Найти $b_1^* = Ab_0$ $\lambda^{(1)} = \max(b_1^*)$

Вычислить первое приближение собственного вектора

$$b_{1(i)} = b_{1(i)}^* / \lambda^{(1)}$$

Алгоритм вычисления собственных векторов

3. Найти $\mathbf{b}_{k+1}^* = \mathbf{A}\mathbf{b}_k$ $\lambda^{(k+1)} = \max(b_{k+1}^*)$

Вычисляем $k+1$ приближение собственного вектора

$$b_{k+1(i)} = b_{k+1(i)}^* / \lambda^{k+1}$$

Вычисляется норма ошибки $\varepsilon = \|\mathbf{b}_{k+1} - \mathbf{b}_k\|$

Полагаем $k=k+1$.

4. Шаг 3 повторяется пока $\varepsilon > \varepsilon_{\text{lim}}$

СОБСТВЕННЫЙ ВЕКТОР

Свойства собственных векторов.

1. Матрица $n \times n$ имеет n собственных векторов.
2. Если собственный вектор умножить на ненулевой коэффициент, то результирующий вектор также является собственным вектором.

СОБСТВЕННЫЙ ВЕКТОР

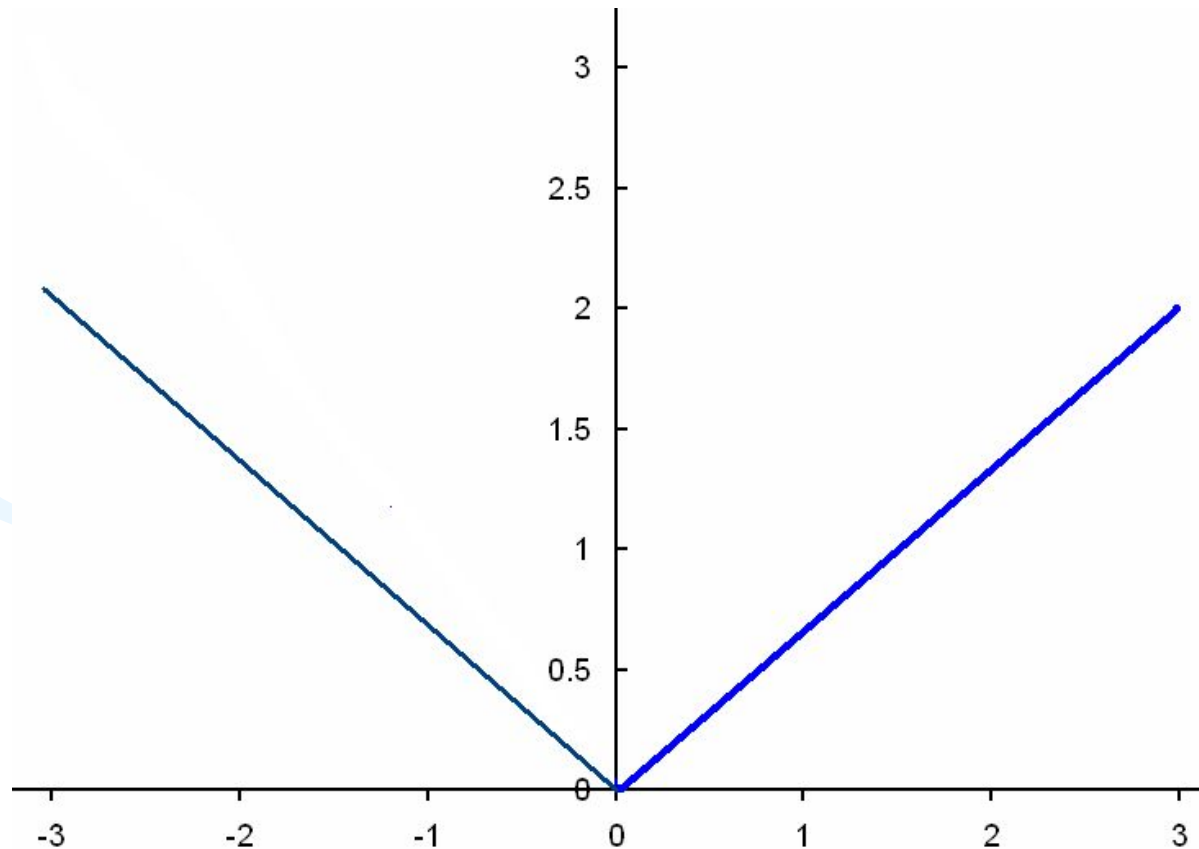
3. Ортогональность собственных векторов

$$b_j^T \mathbf{A} b_i = \begin{cases} a_i & i = j \\ 0 & i \neq j \end{cases}$$

$$b_j^T b_i = \begin{cases} a_i & i = j \\ 0 & i \neq j \end{cases}$$

Собственный вектор

3. Ортогональность собственных векторов



Упражнение. Для следующей квадратной матрицы определить какие векторы являются собственными

$$\begin{bmatrix} 3 & 0 & 1 \\ -4 & 1 & 2 \\ -6 & 0 & -2 \end{bmatrix}$$

Векторы

$$\begin{Bmatrix} 2 \\ 2 \\ -1 \end{Bmatrix} \quad \begin{Bmatrix} -1 \\ 0 \\ 2 \end{Bmatrix} \quad \begin{Bmatrix} -1 \\ 1 \\ 3 \end{Bmatrix} \quad \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix} \quad \begin{Bmatrix} 3 \\ 2 \\ 1 \end{Bmatrix}$$