УДК 539.216:546.824-31

ЛИНЕЙНОЕ ПРОСТОЕ СИММЕТРИЧНОЕ ПРОКРУСТОВО ПРЕОБРАЗОВАНИЕ

В.С. МУХА, А.Н. КУЗЬКОВ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 29 декабря 2015

Разработано линейное преобразование обычных (двухмерных) матриц, которое названо простым симметричным Прокрустовым преобразованием. Предложенное преобразование может быть использовано в распознавании графических представлений физических объектов и в решении иных задач.

Ключевые слова: Прокрустово преобразование, Прокрустово расстояние, регрессионный анализ, линейная регрессия.

Введение

Прокрустовым преобразованием будем называть любое преобразование $f: X \to Y$ $(k \times n)$-матрицы данных X в $(k \times n)$-матрицу Y, обеспечивающее близость результата преобразования (матрицы Y) к некоторой заданной $(k \times n)$-матрице Y_o. Классическим является линейное ортогональное Прокрустово преобразование [1].

Прокрустово преобразование имеет следующую физическую интерпретацию. Набор столбцов каждой из матрицы X, Y, Y_o рассматривается как набор упорядоченных точек, образующих в k-мерном пространстве некоторую фигуру. Прокрустово преобразование матрицы X понимается как преобразование фигуры X в фигуру Y, по возможности более близкую к фигуре Y_o. В связи с такой интерпретацией Прокрустово преобразование можно использовать в распознавании графических представлений физических объектов, а также в решении иных задач.

В связи со сложностью программной реализации классического линейного ортогонального Прокрустова преобразования актуальным является получение иных типов Прокрустова преобразования, приемлемых для решения задач определенных классов. В данной статье разрабатывается преобразование, названное линейным простым симметричным Прокрустовым преобразованием.

Линейная простая симметричная регрессия

Прокрустово преобразование формально может быть получено в рамках теории регрессионного анализа [2]. Пусть (ξ, η) – пара центрированных скалярных случайных величин, и $f(x, y)$ – совместная плотность вероятности этих величин. Случайные величины ξ и η, в предположении их коррелированности, связаны между собой линейной стохастической зависимостью. Поставим задачу выразить эту зависимость с помощью прямой $y = \beta x$ некоторым оптимальным образом: $f(\beta) \to \min_{\beta}$, где $f(\beta)$ – функция (критерий), подлежащая минимизации. Известно, что при выборе критерия оптимальности вида

$$f(\beta) = E((\eta - \beta\xi)^2) \qquad (1)$$

коэффициент β линейной зависимости определяется выражением $\beta = \dfrac{\mu_{\xi\eta}}{\mu_{\xi^2}}$, где

$$\mu_{\xi^2} = E(\xi^2),\ \mu_{\xi\eta} = E(\xi\eta) \qquad (2)$$

теоретические моменты. В этом случае мы получаем функцию $y = \beta x = E(\eta / \xi = x)$, которая является условным математическим ожиданием случайной величины η и обычно называется линейной простой регрессией η на ξ. К такой регрессии мы будем применять термин «линейная простая обычная регрессия» η на ξ. Критерий (1) определяет среднее значение квадрата отклонения случайной величины η от случайной величины $\gamma = \beta\xi$, предсказанной на основании ξ с помощью линейной функции.

Сформулируем иной критерий оптимальности, основанный на отклонении случайной величины η от искомой прямой. Как известно, расстояние от точки (ξ, η) до прямой $Ax + By + C = 0$ определяется выражением [3]: $d = \dfrac{|A\xi + B\eta + C|}{\sqrt{A^2 + B^2}}$.

Следовательно, расстояние от точки (ξ, η) до прямой $y = \alpha + \beta x$ (т.е. прямой $-\beta x + y - \alpha = 0$) будет представлено в виде $d = \dfrac{|\eta - \alpha - \beta\xi|}{\sqrt{\beta^2 + 1}}$, а при $\alpha = 0$ – в виде

$$d = \frac{|\eta - \beta\xi|}{\sqrt{\beta^2 + 1}}. \qquad (3)$$

Выберем критерий оптимальности, равный среднему значению квадрата отклонения d (3) случайной величины η от аппроксимирующей прямой $y = \beta x$:

$$f(\beta) = E(d^2) = \frac{1}{\beta^2 + 1} E((\eta - \beta\xi)^2) = \frac{1}{\beta^2 + 1}(E(\eta^2) - 2\beta E(\xi\eta) + \beta^2 E(\xi^2)). \qquad (4)$$

Если, в дополнение к (2), обозначить $\mu_{\eta^2} = E(\eta^2)$, то функция $f(\beta) = E(d^2)$ (4) примет вид

$$f(\beta) = \frac{1}{\beta^2 + 1}(\mu_{\eta^2} - 2\beta\mu_{\xi\eta} + \beta^2\mu_{\xi^2}). \qquad (5)$$

Функцию $y = \beta x$ с коэффициентом β, полученным при минимизации критерия (4), будем называть линейной простой симметричной (linear simple symmetrical) регрессией η на ξ.

Необходимое условие минимума функции $f(\beta)$ (5) есть уравнение вида

$$\frac{df(\beta)}{d\beta} = \frac{d}{d\beta}\left(\frac{1}{\beta^2 + 1}(\mu_{\eta^2} - 2\beta\mu_{\xi\eta} + \beta^2\mu_{\xi^2})\right) = 0.$$

Выполняя здесь дифференцирование, получим уравнение

$$-2\mu_{\eta^2}\frac{\beta}{(\beta^2 + 1)^2} - 2\mu_{\xi\eta}\frac{1 - \beta^2}{(\beta^2 + 1)^2} + 2\mu_{\xi^2}\frac{\beta}{(\beta^2 + 1)^2} = 0,$$

или уравнение $\mu_{\xi\eta}\beta^2 + (\mu_{\xi^2} - \mu_{\eta^2})\beta - \mu_{\xi\eta} = 0$.

Из этого уравнения имеем

$$\beta = \frac{-(\mu_{\xi^2} - \mu_{\eta^2}) \pm \sqrt{(\mu_{\xi^2} - \mu_{\eta^2})^2 + 4(\mu_{\xi\eta})^2}}{2\mu_{\xi\eta}}. \qquad (6)$$

Минимальное значение критерия (5) определяется путем подстановки (6) в (5). Однако получить простое выражение при этом не удается.

Отметим, что представленная линейная простая симметричная регрессия η на ξ в случае совместного нормального распределения случайных величин ξ, η совпадает с осью эллипса рассеяния (что не так для линейной простой обычной регрессии η на ξ). В связи с этим поле рассеяния двухмерного нормального распределения оказывается симметричным относительно данной прямой, что отражено в названии прямой.

Заметим, что уравнение такой регрессии получено в работе [4] менее прозрачным путем. В [4] эта регрессия названа ортогональной. Представляется, что термин «симметричная» лучше отражает физический смысл данной регрессии.

Линейная векторно-векторная симметричная регрессия

В отличие от предыдущего раздела, рассмотрим не одну, а k пар центрированных скалярных случайных величин (ξ_i, η_i), $i = \overline{1,k}$. В этом случае мы будем иметь k скалярных аппроксимирующих функций $y_i = \beta_i x_i$, $i = \overline{1,k}$, коэффициенты которых определяются выражениями

$$\beta_i = \frac{-(\mu_{\xi_i^2} - \mu_{\eta_i^2}) \pm \sqrt{(\mu_{\xi_i^2} - \mu_{\eta_i^2})^2 + 4(\mu_{\xi_i \eta_i})^2}}{2\mu_{\xi_i \eta_i}}, \tag{7}$$

где $\mu_{\xi_i^2} = E(\xi_i^2)$, $\mu_{\eta_i^2} = E(\eta_i^2)$, $\mu_{\xi_i \eta_i} = E(\xi_i \eta_i)$. (8)

Пусть $\overline{\xi} = (\xi_i)$, $\overline{\eta} = (\eta_i)$, $i = \overline{1,k}$, – два центрированных случайных вектора, и $f(\overline{x}, \overline{y})$, $\overline{x} = (x_i)$, $\overline{y} = (y_i)$, $i = \overline{1,k}$, – их совместная плотность вероятности. Преобразование вида

$$\overline{y} = B\overline{x} \tag{9}$$

с диагональной матрицей преобразования $B = (b_{i,j})$, $i, j = \overline{1,k}$, диагональные элементы которой определяются равенством $b_{i,i} = \beta_i$ с β_i вида (7), назовем векторно-векторной линейной простой симметричной регрессией $\overline{\eta}$ на $\overline{\xi}$. Если векторы $\overline{\xi}$ и $\overline{\eta}$ не являются центрированными, то понятно, что вместо преобразования (9) мы будем иметь преобразование вида $\overline{y} - \overline{\nu}_\eta = B(\overline{x} - \overline{\nu}_\xi)$, то есть аффинное преобразование

$$\overline{y} = \overline{\nu}_\eta + B(\overline{x} - \overline{\nu}_\xi), \tag{10}$$

где, в дополнение к (8),

$$\overline{\nu}_{\overline{\xi}} = E(\overline{\xi}) = (\nu_{\xi_i}),\ \nu_{\xi_i} = E(\xi_i),\ \overline{\nu}_{\overline{\eta}} = E(\overline{\eta}) = (\nu_{\eta_i}),\ \nu_{\eta_i} = E(\eta_i). \tag{11}$$

Линейное простое симметричное Прокрустово преобразование

Пусть теперь $\overline{x}_j$ и $\overline{y}_{o,j}$, $j = \overline{1,n}$, – наблюдения случайных не центрированных векторов $\overline{\xi}$ и $\overline{\eta}$ соответственно, и $f(\overline{x}, \overline{y})$ – их совместная плотность вероятности. Эти наблюдения образуют $(k \times n)$-матрицы наблюдений:

$X = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) = (x_{i,j})$ и $Y_o = (\overline{y}_{o,1}, \overline{y}_{o,2}, ..., \overline{y}_{o,n}) = (y_{o,i,j})$, $i = \overline{1,k}$, $j = \overline{1,n}$.

В этом случае в векторно-векторной аффинной простой симметричной регрессии (10) мы можем теоретические моменты (8), (11) заменить соответствующими эмпирическими и получить тем самым эмпирическую векторно-векторную аффинную простую симметричную регрессию $\overline{y} = \widehat{\overline{\nu}}_\eta + \widehat{B}(\overline{x} - \widehat{\overline{\nu}}_\xi)$, где $\widehat{B} = (\widehat{b}_{i,j})$, $i, j = \overline{1,k}$,

$$\widehat{b}_{i,j} = \begin{cases} \widehat{\beta}_i, & i = j, \\ 0, & i \neq j, \end{cases} \tag{12}$$

$$\widehat{\beta}_i = \frac{-(\widehat{\mu}_{\xi_i^2} - \widehat{\mu}_{\eta_i^2}) \pm \sqrt{(\widehat{\mu}_{\xi_i^2} - \widehat{\mu}_{\eta_i^2})^2 + 4(\widehat{\mu}_{\xi_i\eta_i})^2}}{2\widehat{\mu}_{\xi_i\eta_i}}, \quad (13)$$

$$\widehat{\mu}_{\xi_i\eta_i} = \frac{1}{n}\sum_{j=1}^{n} x_{i,j} y_{o,i,j}, \ \widehat{\mu}_{\xi_i^2} = \frac{1}{n}\sum_{j=1}^{n} x_{i,j}^2, \ \widehat{\mu}_{\eta_i^2} = \frac{1}{n}\sum_{j=1}^{n} y_{o,i,j}^2, \quad (14)$$

$$\widehat{\overline{\nu}}_{\overline{\xi}} = (\widehat{\nu}_{\xi_i}), \ \widehat{\overline{\nu}}_{\overline{\eta}} = (\widehat{\nu}_{\eta_i}), \quad (15)$$

$$\widehat{\nu}_{\xi,i} = \frac{1}{n}\sum_{j=1}^{n} x_{i,j}, \ \widehat{\nu}_{\eta,i} = \frac{1}{n}\sum_{j=1}^{n} y_{o,i,j}, \ i = \overline{1,k}, \ j = \overline{1,n}.$$

Преобразование $Y = \overline{Y} + \widehat{B}(X - \overline{X})$, в котором элементы $\widehat{b}_{i,j}$ матрицы $\widehat{B} = (\widehat{b}_{i,j})$, $i, j = \overline{1,k}$, определяются выражениями (12), (13), (14), а $(k \times n)$-матрицы $\overline{X} = (\widehat{\overline{\nu}}_{\overline{\xi}}, \widehat{\overline{\nu}}_{\overline{\xi}}, ..., \widehat{\overline{\nu}}_{\overline{\xi}})$ и $\overline{Y} = (\widehat{\overline{\nu}}_{\overline{\eta}}, \widehat{\overline{\nu}}_{\overline{\eta}}, ..., \widehat{\overline{\nu}}_{\overline{\eta}})$ состоят из n одинаковых столбцов вида (15), будем называть аффинным (или линейным) простым симметричным Прокрустовым преобразованием матрицы данных X в матрицу данных Y.

В задачах распознавания интерес представляет минимальное значение критерия оптимальности Прокрустова преобразования, которое называется абсолютным Прокрустовым расстоянием между X и Y_o. Оно может быть рассчитано по формуле (5) $d_{\text{абс.}} = \sum_{i=1}^{k} \frac{1}{\widehat{\beta}_i^2 + 1}(\widehat{\mu}_{\eta_i^2} - 2\widehat{\beta}_i\widehat{\mu}_{\xi_i\eta_i} + \widehat{\beta}_i^2\widehat{\mu}_{\xi_i^2})$.

Однако поскольку выбранный критерий оптимальности $f(\beta)$ (5) не ориентирован непосредственно на близость матриц Y и Y_o, то рекомендуется руководствоваться критерием непосредственной близости этих матриц и рассчитывать Прокрустово расстояние по формуле

$$d_{\text{абс.}} = \| Y_o - Y \|_E^2 = \sum_{i=1}^{k}\sum_{j=1}^{n} (y_{o,i,j} - y_{i,j})^2 = tr\left((Y_o - Y)(Y_o - Y)^T\right), \quad (16)$$

где $\| Y_o - Y \|_E^2$ – квадрат Евклидовой норма матрицы $Y_o - \overline{Y}$. На практике для достижения инвариантности к размерам распознаваемых образов целесообразно использовать относительное Прокрустово расстояние $d_{\text{отн.}} = d_{\text{абс.}} / \| Y_o - \overline{Y} \|_E^2$, где

$$\| Y_o - \overline{Y} \|_E^2 = \sum_{i=1}^{k}\sum_{j=1}^{n} (y_{o,i,j} - \overline{y}_{i,j})^2 = tr\left((Y_o - \overline{Y})(Y_o - \overline{Y})^T\right).$$

Необходимо также выбрать комбинацию знаков «плюс» или «минус» в формуле (13) для всех $i = \overline{1,k}$, обеспечивающую минимум Прокрустову расстоянию (16).

LINEAR SIMPLE SYMMETRICAL PROCRUSTES TRANSFORMATION

V.S. MUKHA, A.N. KUZKOU

Abstract

It is designed the Linear transformation of usual (two-dimensional) matrixes, which is named as simple symmetrical Procrustes transformation. The offered transformation can be used in recognition of the graphic presentations of physical objects and in decision of other problems.

Список литературы

1. *Crosilla F.* // Technical report. Part 1. 1999. P. 69–78.
2. *Муха В.С.* // Кибернетика и системный анализ. 2007. № 3. С. 138–143.
3. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. М., 1973.
4. *Крамер Г.* Математические методы статистики. М., 1975.