1 Билинейные функции

Определение 1.1. *Пусть V линейное пространство над полем K. Отображение $V \times V \xrightarrow{f} K$ называется билинейной функцией, если $\forall\ x, y, z \in V, \forall\ \alpha \in K$ справедливы равенства*

1. $f(x+y, z) = f(x, z) + f(y, z)$
2. $f(x, y+z) = f(x, y) + f(x, z)$
3. $f(\alpha x, y) = \alpha f(x, y)$
4. $f(x, \alpha y) = \alpha f(x, y)$

Билинейная функция f называется симметричной, если $f(x, y) = f(y, x)$.

Пример билинейной функции:[1]

$$f(x, y) = x^\top A y,\ \text{где } A \in \mathbb{R}^{n\times n}, x \in \mathbb{R}^n, y \in \mathbb{R}^n \tag{1}$$

Выражение билинейной функции через координаты векторов.

Пусть $e = \{e_1, .., e_n\}$ – базис, $x = x_1 e_1 + ... + x_n e_n,\ y = y_1 e_1 + ... + y_n e_n$. Из определения билинейной функции следует, что

$$f(x, y) = \sum_{i,j} f(e_i, e_j) x_i y_j. \tag{2}$$

Полагая $x_e = (x_1, .., x_n)^\top, y_e = (y_1, .., y_n)^\top, F_e = (f(e_i, e_j))$, можно записать равенство (2) в матричном виде

$$f(x, y) = x_e^\top F_e y_e. \tag{3}$$

Определение 1.2. *F_e называется матрицей билинейной функции в базисе e.*

Изменение матрицы билинейной функции при изменении базиса.

Пусть F_e и $F_{e'}$ – матрицы билинейной функции $f(x, y)$ в базисах e и e' соответственно, P – матрица перехода от e к e', тогда

$$x_e = P x_{e'}, y_e = P y_{e'} \text{ и } f(x, y) = x_e^\top F_e y_e = x_{e'}^\top P^\top F_e P y_{e'}$$

Следовательно,

$$F_{e'} = P^\top F_e P. \tag{4}$$

Определение 1.3. *Матрицы A, B называются конгруентными, если существует невырожденная матрица P такая, что $A = P^\top B P$. Разные конгруентные матрицы представляют одну и ту же билинейную функцию в разных базисах.*

Определение 1.4. *Правая часть равенств (2) и (3) называется билинейной формой. F_e называется матрицей билинейной формы.*

[1] При некоторых ограничениях на матрицу A правая часть равенства может задавать скалярное произведение векторов

Определение 1.5. *Если $f(x,y)$ билинейная функция, то функция $g(x) = f(x,x)$ называется квадратичной. а выражение $\sum_{i,j} f_{ij}x_ix_j$ называется квадратичной формой.*

Замечание 1. *Равенства* (3) *и* (5) *означают, что алгебраические свойства билинейной и квадратичной функции полностью зависят от их матриц.*

Если F – матрица билинейной формы $f(x,y)$, то квадратичную форму $f(x,x)$ можно записать с помощью симметричной матрицы $G = (F + F^\top)/2$, а именно:

$$g(x) = x^\top Gx. \tag{5}$$

Очевидно, что матрицы симметричной билинейной формы $f(x,y)$ и квадратичной формы $f(x,x)$ равны. Следовательно, "с матричной точки зрения" эти формы неразличимы.

Далее изложены результаты, относящиеся к квадратичным формам. Будут обсуждаться три основных вопроса:

Вопрос 1. Для всякой ли квадратичной формы $f(x)$ существует замена $x = Py$[1] такая, что форма

$$g(y) = f(Py) = d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2, \tag{6}$$

не содержит произведений неизвестных?[2]

Вопрос 2. Как найти замену $x = Py$ приводящую форму $f(x)$ к канонической форме.

Вопрос 3. Задана матрица $A \in \mathbb{R}^{n\times n}$. Верно ли, что для всякого $x \in \mathbb{R}^n, x \neq 0$ справедливо неравенство $x^\top Ax > 0$.[3]

Приведение квадратичной формы к каноническому виду

Ранее, в разделе "Линейные преобразования" было доказано, что для любой симметричной матрицы $A \in \mathbb{R}^{n\times n}$ существует ортогональная матрица P, составленная из собственных векторов матрицы A и диагональная матрица $D = diag(\lambda_1, \ldots, \lambda_n)$ с собственными числами матрицы A, для которых справедливо равенство

$$D = P^{-1}AP = P^\top AP.$$

Следовательно, квадратичная форма $x^\top Ax$ заменой $x = Py$ приводится к канонической форме (6), где $d_1, d_2, \ldots, d_n$ – это собственные числа матрицы A.

Таким образом, есть положительный ответ на Вопрос 1 и один из возможных ответов на Вопрос 2. Рассмотрим еще три метода приведения к каноническому виду

[1] Иначе говоря: существует ли базис $P = (p_1, p_2, \ldots, p_n)$?

[2] Форма $g(y)$ называется *канонической* для $f(x)$.

[3] Если неравенство верно, то матрица A и квадратичная форма $x^\top Ax$ называются *положительно определенными.*

Метод выделения полного квадрата(метод Лагранжа).

Суть метода полностью раскрывается следующими двумя примерами.

Пример 1.1. Привести к каноническому виду $f(x_1, x_2, x_3) = x_1x_2 + \underline{6x_1x_3} + \underline{2x_2x_3} + \underline{x_3^2}$.

Решение. Форма содержит квадрат одной из переменных, а именно, x_3^2. Вот относительно x_3 и будем выделять полный квадрат, т.е. найдем форму $(t_1x_1 + t_2x_2 + x_3)^2$, содержащую подчернутые слагаемые.[1]

$$6x_1x_3 + 2x_2x_3 + x_3^2 = (\underbrace{3x_1 + x_2 + x_3}_{y_3})^2 - 9x_1^2 - x_2^2 - 6x_1x_2 = \tag{7}$$

$$= -9x_1^2 - x_2^2 - 6x_1x_2 + y_3^2.$$

Подставляем в f

$$f(x_1, x_2, x_3) = \underbrace{-9x_1^2 - x_2^2 - 6x_1x_2}_{g(x_1, x_2)} + y_3^2 = g(x_1, x_2) + y_3^2 \tag{8}$$

Аналогичные преобразования выполняем в форме g. Вычисления заканчиваются, когда будет получена форма, не содержащая произведений переменных.⧫

Пример 1.2. Привести к каноническому виду

$$f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Решение. Форма не содержит квадратов. Это следует исправить стандартной заменой

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases} \tag{9}$$

$$f(x_1, x_2, x_3) = g(y_1, y_2, y_3) = 2y_1^2 - 2y_2^2 + 2y_2y_3.$$

Форма g содержит квадраты неизвестных, следовательно, к ней примени́м прием, описанный выше.⧫

[1] $(a + b + ... + d + c)^2 = a^2 + b^2 + ... + d^2 + c^2 + 2ab + ... + 2ad + 2ac + 2bd + 2bc$

Матричный метод. Как и в методе Лагранжа поэтапно избавляемся от произведений переменных, но делаем это, зануляя элементы матрицы квадратичной формы вне главной диагонали. Вернемся к примеру 1. Выпишем матрицы исходной квадратичной формы и формы (8)

$$A = \begin{bmatrix} 0 & 1/2 & 3 \\ 1/2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} B = \begin{bmatrix} -9 & -5/2 & 0 \\ -5/2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

Сопоставляя элементы матриц, можно заметить, что B получается из A, если выполнить преобразования строк и столбцов:

$$(1) - 3(3), (2) - (3), \tag{11}$$

$$[1] - 3[3], [2] - [3]. \tag{12}$$

Поскольку элементарные преобразования столбцов можно заменить умножением A справа на матрицу[1]

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & -1 & 1 \end{bmatrix},$$

а элементарные преобразования строк на матрицу

$$P^\top = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix},$$

то $B = P^\top AP$, матрицы A и B конгруентны.

Отсюда следует вывод: для приведения квадратичной формы к каноническому виду, достаточно *согласованными*[2] элементарными преобразованиями строк и столбцов привести матрицу формы к диагональной матрице, придерживаясь двух правил:

1. Ведущие элементы следует выбирать только на главной диагонали подматрицы, преобразуемой на очередном этапе.

2. Если все диагональные элементы равны нулю, то выбрать строку i и столбец j, на пересечении которых расположен ненулевой элемент, и выполнить согласованные преобразования $[i]+[j]$ и $(i)+(j)$. В новой подматрице будет ненулевой диагональный элемент.

Следовательно, замена $x = Py$ приводит $x^\top Ax$ к канонической форме.

Метод сопряженных направлений.

Определение 1.6. *Векторы $p, q \in \mathbb{R}^n$ называются сопряженными относительно симметричной матрицы $A \in \mathbb{R}^{n\times n}$, если $p^\top Aq = 0$.*

---

[1]Вспомните о связи элементарных преобразований строк и столбцов с умножением матриц

[2]каждое преобразование строк выполняется и над столбцами с теми же номерами

Этап 1. Выберем h такой, что $h^\top Ah \neq 0$. Полагаем

$$p_1 = h,\ L_1 = \{x | p_1^\top Ax = 0\}.$$

Предположим, что уже найдены векторы $p_1, .., p_{k-1}$ и подпространство L_{k-1}.

Этап k. Рассмотрим 2 возможных случая.

а) Существует $h \in L_{k-1}$ такой, что $h^\top Ah \neq 0$. Полагаем

$$p_k = h,\ L_k = \{x \in L_{k-1} | p_k^\top Ax = 0\}.$$

Если $k < n$, то переходим к следующему этапу.

b) $\forall x \in L_{k-1}$ верно равенство $x^\top Ax = 0$. Вычисления заканчиваются. Каноническим базисом будет объединение системы $p_1, .., p_{k-1}$ и любого базиса $h_k, .., h_n$ подпространства L_{k-1}.

Замена $x = p_1y_1 + \ldots + p_{k-1}y_{k-1} + h_ky_k + \ldots + h_ny_n$ приводит квадратичную форму $x^\top Ax$, к каноническому виду

$$p_1^\top Ap_1y_1^2 + \cdots + p_{k-1}^\top Ap_{k-1}y_{k-1}^2 + h_k^\top Ah_ky_k^2 + \cdots + h_n^\top Ah_ny_n^2.$$

Обоснование выполним в предположении, что на этапе k имел место случай b). Другой вариант рассматривается аналогично.

(1) Линейная независимость векторов $p_1, .., p_{k-1}, h_k, .., h_n$.

$$\alpha_1p_1 + .. + \alpha_{k-1}p_{k-1} + \alpha_kh_k + .. + \alpha_nh_n = 0 \to \tag{13}$$

$$p_i^\top A(\alpha_1p_1 + .. + \alpha_{k-1}p_{k-1} + \alpha_kh_k + .. + \alpha_nh_n) = 0 \to$$

$$\alpha_ip_iAp_i = 0 \to \alpha_i = 0\, \forall i \leq k-1.$$

Уравнение (13) превращается в уравнение $\alpha_kh_k + .. + \alpha_nh_n = 0$, которое имеет только нулевое решение.

(2) Соотношения $p_i^\top Ap_j = 0 (i \neq j)$ и $p_i^\top Ah_j = 0$ очевидны. Далее,

$$(h_i + h_j)^\top A(h_i + h_j) = 0 \to h_i \top Ah_i + 2h_i^\top Ah_j + h_j \top Ah_j \to h_i^\top Ah_j = 0.$$

Следовательно, $p_1, .., p_{k-1}, h_k, .., h_n$ – канонический базис.

Пусть P, Q произвольные матрицы и существует произведение PQ. Столбцы матрицы PQ принадлежат линейной оболочке столбцов матрицы P, следовательно, $\operatorname{rank} PQ \leq \operatorname{rank} P$. Если, дополнительно, Q обратима, то $\operatorname{rank} PQ = \operatorname{rank} P$. Аналогично доказывается, что $\operatorname{rank} QP = \operatorname{rank} P$, если Q обратима. Это означает, что ранги конгруентных матриц равны между собой.

Теорема инерции, положительные формы

Определение 1.7. *Рангом квадратичной функции называется ранг ее матрицы.*

Утверждение 1.1. (Теорема инерции.) Число положительных коэффициентов в канонической форме квадратичной функции f независит от выбора канонического базиса.

Д о к а з а т е л ь с т в о. Рассмотрим конгруентные формы ранга r
$f(x_1,..,x_n)=\alpha_1 x_1^2+..+\alpha_r x_r^2$ и $g(y_1,..,y_n)=\beta_1 y_1^2+..+\beta_r y_r^2$.
Пусть переменные связаны равенствами $x_i=\sum_j a_{ij}y_j$ $(i=1,..,n)$;

$$\alpha_1,..,\alpha_k>0,\ \alpha_{k+1},..,\alpha_r<0,$$

$$\beta_1,..,\beta_s>0,\ \beta_{s+1},..,\beta_r<0.$$

Требуется доказать, что $k=s$.

Предположим противное, пусть, для определенности, $k<s$. Система уравнений

$$\begin{cases}\sum_j a_{ij}y_j=0\ (i=1,..,k),\\ y_j=0\ (j=s+1,..,n)\end{cases}$$

состоит из $n-s+k<n$ уравнений, поэтому имеет ненулевое решение $(y_1^0,..,y_n^0)$. Положим $x_i^0=\sum_j a_{ij}y_j^0$ $(i=1,..,n)$. Так как $y_j^0\neq 0$ хотя бы для одного $j\le s$, то

$$g(y_1^0,..,y_n^0)=\beta_1(y_1^0)^2+..+\beta_s(y_s^0)^2>0.$$

С другой стороны,

$$g(y_1^0,..,y_n^0)=f(x_1^0,..,x_n^0)=\alpha_1(x_{k+1}^0)^2+..+\alpha_r(x_r^0)^2\le 0.$$

Полученное противоречие доказывает утверждение. $\square$

Определение 1.8. *Угловым минором k - го порядка матрицы A называется определитель подматрицы* $A\begin{pmatrix}1&2&...&k\\1&2&...&k\end{pmatrix}$.

Утверждение 1.2. (Формула Якоби.) Пусть $d_1,d_2,..d_n$ - угловые миноры симметричной матрицы A, причем $d_i\neq 0$, при $i\le m$ и $d_i=0$, при $i>m$. Тогда A конгруентна матрице

$$D=diag(d_1,\frac{d_2}{d_1},...,\frac{d_m}{d_{m-1}},0,\dots,0).$$

Д о к а з а т е л ь с т в о. Применим к матрице матричный метод приведения к диагональной форме. Так как $a_{1,1}=d_1\neq 0$, то можно выполнить согласованные преобразования с ведущим элементом $a_{1,1}$ и получить матрицу

$$A'=\begin{bmatrix}a_{1,1}&0&0&..&0\\0&a'_{2,2}&a'_{2,3}&...&a'_{2,n}\\0&a'_{3,2}&a'_{3,3}&...&a'_{3,n}\\..&..&..&..&..\\0&a'_{n,2}&a'_{n,3}&...&a'_{n,n}\end{bmatrix}$$

Угловые миноры при этом не меняются, следовательно, угловые миноры матрицы $A'\begin{pmatrix}2&3&...&n\\2&3&...&n\end{pmatrix}$ равны $d_2/d_1,..d_m/d_1,0,.,0$. В частности, $a'_{2,2}=d_2/d_1\neq 0$, поэтому $a'_{2,2}$ – ведущий элемент на втором этапе. Угловые миноры матрицы $A''\begin{pmatrix}3&4&...&n\\3&4&...&n\end{pmatrix}$

равны $d_3/d_2, ..d_m/d_2, 0, ., 0$ и т.д.. $\square$

Определение 1.9. *Квадратичная форма $f(x)$ называется положительно определенной*[1]*, если $\forall x \neq 0$ $f(x) > 0$.*

Аналогично вводятся понятия других *знакоопределенных* форм : неотрицательно, неположительно и отрицательно определенной. Каждая из перечисленных форм называется также знакоопределенной.

Утверждение 1.3. (Критерий Сильвестра.) Квадратичная форма $f(x) = x^\top Ax$ положительно определена тогда и только тогда, когда угловые миноры $d_1, .., d_n$матрицы A положительны.

Доказательство. Достаточность. Согласно формуле Якоби форма $f(x)$ конгруентна форме

$$g(y) = d_1 y_1^2 + .. + \frac{d_k}{d_{k-1}} y_k^2 + .. + \frac{d_n}{d_{n-1}} y_n^2. \tag{14}$$

Коэффициенты формы $g(y)$ положительны, поэтому она положительна, следовательно, $f(x)$ также положительна.

Необходимость. Пусть $f(x)$ положительна.

Покажем сначала, что $d_i \neq 0 \, \forall i$. Предположим обратное, пусть $d_k = 0$ для некоторого k. Рассмотрим матрицу$B = A \begin{pmatrix} 1 & 2 & ... & k \\ 1 & 2 & ... & k \end{pmatrix}$. Так как $\det B = 0$, то $\exists z = (z_1 ... z_k)^\top \neq 0$ такой, что $Bz = 0$ и $z^\top Bz = 0$. Рассмотрим n - мерный вектор $z' = (z_1 ... z_k \, 0..0)^\top \neq 0$. Так как $f(z') = z'^\top Az' = z^\top Bz = 0$, то получено противоречие с предположением о положительной определенности $f(x)$. Таким образом, все угловые миноры отличны от нуля.

Предположим теперь, что среди $d_1, .., d_n$ есть отрицательные. Пусть k - наименьший номер, для которого $d_k < 0$. Подставляя в (14) значения $y'_k = 1, y'_i = 0 \, \forall i \neq k$ получаем неравенство $g(y') < 0$, противоречащее положительной определенности $g(y)$. $\square$

Приведение пары форм.*

Рассматривается пара форм $f = x^\top Ax$, $g = x^\top Bx$, причем f положительно определена. Оказывается можно Найти преобразование переменных $x = Sy$, приводящее f к нормальному виду, а g – к каноническому.

Метод состоит в следующем:

1. Найти матрицу P такую, что $P^\top AP = E$. После замены $x = Pz$ формы преобразуются к виду $f(z) = z^\top z$ и $g(z) = z^\top B'z$, где $B' = P^\top BP$.

[1]Или просто положительной.

2. Матрица B' симметрична, потому ее собственные числа $\lambda_1, .., \lambda_n$ вещественны и она имеет ортонормированный набор $h_1, .., h_n$ собственных векторов.Образуем матрицу $H = (h_1, .., h_n)$ и делаем замену $z = Hy$. Так как

$H^\top H = E$ и $H^\top B' H = diag(\lambda_1, .., \lambda_n)$, то$f(y) = y^\top H^\top H y = y^\top y = y_1^2 + ... + y_n^2$

При вычислениях могут оказаться полезными следующие замечания.

Замечание 2. *Собственные числа матрицы B' удовлетворяют уравнению $det(B - \lambda A) = 0$*

Замечание 3. *Для каждого $i \in \{1, .., n\}$ i - й столбец s_iудовлетворяет уравнению*

$$(B - \lambda_i A)x = 0. \tag{15}$$

Замечание 4. *Предположим, что все собственные числа различны.Пусть q_i решение уравнения (15), $Q = (q_1..q_n)$.Замена $x = Qy$ приводит обе формы f, g к каноническому виду. Действительно, из равенств*

$$\lambda_i q_j^\top A q_i = q_j^\top (\lambda_i A q_i) = q_j^\top B q_i = q_i^\top B q_j = q_i^\top \lambda_j A q_j = \lambda_j q_i^\top A q_j = \lambda_j q_j^\top A q_i,$$

при $\lambda_i \neq \lambda_j$ следует, что $q_j^\top A q_i = q_j^\top B q_i = 0$, при $i \neq j$, поэтому матрицы $Q^\top A Q$ и $Q^\top B Q$ диагональны.

Замечание 5. *Если не все собственные числа различны, то к базису множества решений уравнения (15) надо применить процесс ортогонализации, причем в качестве матрицы скалярного произведения используется матрица A.*