

Сборник заметок по линейной алгебре и сопряженным вопросам

Подвойский А.О.

Содержание

1 Ранг матрицы	1
2 Базис линейного пространства. Координаты векторов	2
3 Мера обусловленности матрицы	3
4 Линейно зависимые и линейно независимые системы	3
4.1 Свойства линейно зависимых и линейно независимых столбцов	4
5 Система m линейных алгебраических уравнений с n неизвестными	4
6 Теорема (правило) Крамера	5
7 Условие совместности системы линейных уравнений. Теорема Кронекера-Капелли	5
8 Общее решение системы линейных алгебраических уравнений	6
9 Решение систем уравнений с помощью полуобратных матриц	6
10 Псевдорешения системы линейных уравнений	8
11 Свойства решений однородной системы	8
12 Функциональные матрицы скалярного аргумента	9
13 Производные скалярной функции по векторному аргументу	9
14 Производные от векторной функции векторного аргумента	11
15 Правила дифференцирования по векторному аргументу	12
16 Производные матричной функции по векторному аргументу	13
17 Линейные и квадратичные формы	14
Список литературы	15

1. Ранг матрицы

Базисный минор — это отличный от нуля минор наивысшего порядка.

Ранг матрицы это наибольший порядок отличного от нуля минора этой матрицы. Отличные от нуля миноры наивысшего порядка называют *базисными*. Столбцы матрицы, на которых располагается хотя бы один базисный минор этой матрицы *линейно независимы* и называются *базисными столбцами*.

Если ранг матрицы совпадает с числом ее столбцов $\text{rg } A_{m \times n} = n$, то все столбцы этой матрицы линейно независимы [4, стр. 31]. Существенным является то, что ранг матрицы не изменяется при элементарных преобразованиях [4, стр. 34].

Ранг матрицы можно вычислить методом окаймляющих миноров [1], [4, стр. 31].

2. Базис линейного пространства. Координаты векторов

Всякую систему векторов линейного пространства X называют *базисом* или *базой этого пространства*, если эта система векторов *линейно независима* и любой вектор пространства X линейно выражается через векторы этой системы.

Если базис пространства конечен, т.е. состоит из конечного числа векторов, то он представляет собой конечную максимальную линейно независимую систему векторов пространства, и обратно, *любая конечная максимальная линейно независимая система векторов* линейного пространства является *базисом* этого пространства.

Линейное пространство X называют конечномерным, если оно обладает хотя бы одним базисом, состоящим из конечного числа векторов. Конечномерное пространство может обладать многими различными базисами. Число векторов в каждом базисе конечномерного пространства одинаково. Это число называют *размерностью* пространства и обозначают $\dim X = n$. Пространство X при этом называют n -мерным.

То есть размерность пространства $\dim X$ – это число векторов в базисе пространства.

Пусть линейное пространство X_n обладает базисом

$$e : e_1, e_2, \dots, e_n.$$

Тогда любой вектор x из X_n единственным образом представляется в виде

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n = e \cdot [x]_e.$$

Числа x_1, x_2, \dots в разложении называют координатами вектора x в базисе и записывают $x (x_1, x_2, \dots, x_n)_e^T$.

Система линейных уравнений, имеющая *хотя бы одно* решение, называется *совместной*. Если у системы *нет решений*, то она называется *несовместной* [4, стр. 41].

Система совместна тогда и только тогда, когда ранг матрицы A системы равен рангу расширенной матрицы $A|b$: $\text{rg } A = \text{rg}(A|b)$. Ранг матрицы определяется порядком базисного минора.

Если ранг матрицы совпадает с числом неизвестных и система совместна, то она имеет строго одно решение. Если же ранг матрицы меньше числа неизвестных и матрица совместна, то она имеет бесконечное множество решений.

Система называется *однородной*, если все свободные члены ее уравнения равны нулю. Кроме того, *однородная система всегда совместна*, так как она обладает, по крайней мере, нулевым решением $x_1 = x_2 = \dots = x_n = 0$.

Для того чтобы система линейных однородных уравнений имела *только нулевое* решение необходимо и достаточно, чтобы ранг матрицы совпадал с числом неизвестных системы [4, стр. 43]. В частности, однородная система n линейных уравнений с n неизвестными имеет только нулевое решение, если ее определитель отличен от нуля. Если определитель однородной системы отличен от нуля, то это значит, что ранг матрицы совпадает с числом столбцов (неизвестных), а однородная система всегда совместна. В этом случае может существовать только одно решение – нулевое.

Для того чтобы однородная система имела ненулевые решения, необходимо и достаточно, чтобы ранг ее матрицы был меньше числа неизвестных системы, $\text{rg} A < \#x_k$. Любая однородная система, в которой число уравнений меньше числа неизвестных имеет ненулевые решения.

3. Мера обусловленности матрицы

Мера (или число) обусловленности матрицы A определяется как [3, стр. 306]

$$\nu(A) = \|A\| \|A^{-1}\|$$

Поскольку *любая норма* матрицы не меньше своего наибольшего по модулю собственного значения, то $\|A\| \geq \max |\lambda_A|$ и поскольку собственные значения матриц A и A^{-1} взаимно обратны, то

$$\|A^{-1}\| \geq \max \frac{1}{|\lambda_A|} = \frac{1}{\min |\lambda_A|}$$

Таким образом, мера обусловленности матрицы A

$$\boxed{\nu(A) \geq \frac{\max |\lambda_A|}{\min |\lambda_A|} \geq 1}$$

В частности, при $A = A^T$ (то есть если матрица симметричная) имеем $\|A\|_2 = \max |\lambda_A|$.

Следовательно, в случае нормы $\|\cdot\|_2$

$$\nu(A) = \frac{\max |\lambda_A|}{\min |\lambda_A|}.$$

4. Линейно зависимые и линейно независимые системы

Система из k столбцов A_1, \dots, A_k называется *линейно зависимой*, если существуют такие числа $\alpha_1, \dots, \alpha_k$ не все равные нулю одновременно, что [1, стр. 128]

$$\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_k A_k = o,$$

где o – нулевой вектор соответствующего размера.

То есть другими словами система k столбцов называется *линейно зависимой*, если эти столбцы *суммируются в нулевой столбец* для нетривиального случая коэффициентов α_i (когда эти коэффициенты не все одновременно равны нулю).

Система из k столбцов называется *линейно независимой*, если $\sum_{j=1}^k \alpha_j A_j = o$ возможно только в тривиальном случае, т.е. когда $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$.

4.1. Свойства линейно зависимых и линейно независимых столбцов

Замечание

Понятия линейной зависимости и линейной независимости формулируются одинаково как для строк, и так для столбцов

- Если в систему входит *нулевой столбец*, то она *линейно зависима*,
- Если в систему входит *два равных столбца*, то она *линейно зависима*,
- Если в системе столбцов имеется два пропорциональных столбца $A_i = \lambda A_j$, то она линейно зависима,
- Любые столбцы, входящие в *линейно независимую систему*, образуют *линейно независимую подсистему*,
- Система столбцов, содержащая *линейно зависимую подсистему*, сама *линейно зависима*,
- Если система столбцов A_1, \dots, A_k – линейно независима, а после присоединения к ней столбца A – оказывается линейно зависимой, то столбец A можно разложить по столбцам A_1, \dots, A_k и притом единственным образом, т.е. коэффициенты определяются однозначно.

5. Система m линейных алгебраических уравнений с n неизвестными

Матричная запись неоднородной системы уравнений имеет вид

$$Ax = b,$$

а однородной

$$Ax = o,$$

где o в правой части обозначает нулевой столбец размеров $m \times 1$.

Эту матричную запись неоднородной системы уравнений можно представить в эквивалентной форме

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Тогда решение системы представляется столбцом

$$x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

и удовлетворяете равенству

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \alpha_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \alpha_2 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \alpha_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

т.е. столбец свободных членов b является линейной комбинацией столбцов матрицы системы.

6. Теорема (правило) Крамера

Система называется **совместной**, если она имеет *хотя бы одно решение*. Система называется **несовместной**, если она *не имеет ни одного решения*.

Если определитель $\Delta = \det A$ матрицы системы n линейных независимых уравнений с n неизвестными отличен от нуля ($\det A \neq 0$), то система имеет *единственное* решение, которое находится по формулам

$$x_i = \frac{\Delta_i}{\Delta}, \quad i = 1, \dots, n, \quad (\Delta = \det A \neq 0),$$

где Δ_i – определитель матрицы, полученной из матрицы системы $A = [a_{ij}]_{i,j=1}^n$ заменой i -ого столбца столбцом свободных членов.

ЗАМЕЧАНИЕ: на практике при больших n правило Крамера не применяется!

Если $\Delta = 0$ (матрица коэффициентов системы вырождена) и хотя бы один определитель $\Delta_i \neq 0$, то система *несовместна*, т.е. не имеет ни одного решения. Если же $\Delta = \Delta_1 = \Delta_2 = \dots, \Delta_n = 0$, то возможны два случая: либо система несовместна (не имеет ни одного решения), либо система имеет бесконечно много решений [1, стр. 188].

7. Условие совместности системы линейных уравнений. Теорема Кронекера-Капелли

Рассмотрим систему m линейных уравнений с n неизвестными. Составим блочную матрицу, приписав к матрице A справа столбец свободных членов b . Получим *расширенную матрицу системы*

$$(A | b)_{m \times (n+1)} = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

Эта матрица содержит всю информацию о системе уравнений, за исключением обозначений неизвестных.

Теорема Кронекера-Капелли. Система $Ax = b$ *совместна* (т.е. имеет хотя бы одно решение) тогда и только тогда, когда ранг матрицы системы равен рангу расширенной матрицы $\text{rg } A = \text{rg}(A | b)$.

Если $\text{rg } A \neq \text{rg}(A | b)$, то система несовместна – не имеет решений.

Если система имеет решение, то столбец свободных членов есть линейная комбинация столбцов матрицы системы. Поэтому при вычеркивании столбца b из расширенной матрицы $(A \mid b)$ ее ранг не изменяется. Следовательно, $\text{rg}(A \mid b) = \text{rg } A$.

ЗАМЕЧАНИЕ: теорема Кронекера-Капелли дает лишь критерий существования решения системы, но не указывает способа отыскать этого решения.

8. Общее решение системы линейных алгебраических уравнений

Неизвестные, которым соответствуют столбцы, входящие в базисный минор, называются *базисными переменными*, остальные неизвестные – *свободными переменными*.

Общее решение системы, выражающее базисные переменные через свободные, имеет вид [1, стр. 192]

$$\begin{cases} x_1 = b'_1 - a'_{1,r+1}x_{r+1} - \dots - a'_{1,n}x_n, \\ \dots \\ x_r = b'_r - a'_{r,r+1}x_{r+1} - \dots - a'_{r,n}x_n, \end{cases}$$

где x_1, x_2, \dots, x_r – базисные переменные; $x_{r+1}, x_{r+2}, \dots, x_n$ – свободные переменные.

Частное решение системы – решение системы, получающееся из общего решения, заданием конкретных значений свободными переменным.

Пусть x – решение неоднородной системы. Тогда любое решение x неоднородной системы можно представить в виде $x = x^{\text{н}} + x^{\text{о}}$, где $x^{\text{о}}$ – решение однородной системы.

Говорят, что *общее решение* неоднородной системы есть сумма *частного решения* неоднородной системы и *общего решения* соответствующей однородной системы [1, стр. 200]

$$x = x^{\text{н}} + C_1\varphi_1 + C_2\varphi_2 + \dots + C_{n-r}\varphi_{n-r}.$$

9. Решение систем уравнений с помощью полуобратных матриц

Требуется решить систему линейных уравнений

$$Ax = b,$$

где A – произвольная матрица размера $m \times n$.

Если матрица системы нулевая $A = O$, то система либо несовместна (при $b = o$), либо имеет бесконечное множество решений (при $b = o$ любой подходящий по размерам столбец x является решением). Далее рассматривается случай ненулевой матрицы A .

Пусть A^{-1} – матрица, полуобратная к матрице системы A . Используя определение полуобратной матрицы, неоднородную систему $Ax = b$ можно переписать так

$$AA^{-1}Ax = b.$$

Если x – решение системы, то подставляя $Ax = b$ в левую часть последнего соотношения

$$AA^{-1}Ax = b, \quad \rightarrow \quad AA^{-1}b = b.$$

Тогда

$$(E_m - AA^{-1})b = o.$$

Это необходимое и достаточное условие совместности системы.

Решением системы будет $x = A^{-1}b$. Но поскольку *полуобратная матрица* определена *неоднозначно*, то эта формула фактически задает множество решений системы. Преобразуем так, чтобы была видна структура этого множества, в частности, выявим количество независимых параметров

$$A_0^{-1} = T\Lambda^T S = T \left(\begin{array}{c|c} E_r & O \\ \hline O & O \end{array} \right) S,$$

где S и T – элементарные матрицы порядков n и m соответственно, Λ – матрица простейшего вида, эквивалентная матрице A ($\Lambda \sim A$), $\text{rg } A$.

Теорема о совместности неоднородной системы и о структуре ее общего решения. Неоднородная система $Ax = b$ совместна тогда и только тогда, когда столбец свободных членов является решением однородной системы $\Psi b = o$. Если система $Ax = b$ совместна, то ее общее решение имеет вид [1, стр. 205]

$$x = x^h + x^o = A_0^{-1} b + \Psi c = T \left(\begin{array}{c|c} E_r & O \\ \hline O & O \end{array} \right) S b + T \left(\begin{array}{c} O \\ \hline E_{m-r} \end{array} \right) c, \quad \Psi = \left(O \mid E_{m-r} S \right),$$

где T, S – элементарные преобразующие матрицы, $c = (C_1 \dots C_{n-r})^T$ – столбец произвольных постоянных.

Алгоритм применения полуобратной матрицы:

1. Привести матрицу A системы $Ax = b$ к простейшему виду: $\Lambda = SAT$. При этом находятся элементарные преобразующие матрицы S и T , а также ранг $r = \text{rg } A \geq 1$.
2. Проверить условие совместности системы $\Psi b = o$. При $r = m$ система совместна. Если $r < m$, то составить матрицу $\Psi = (O \mid E_{m-r}) S$ и проверить условие $\Psi b = o$. Если условие выполняется, то система совместна. В противном случае система несовместна и процесс решения заканчивается.
3. Найти частное решение неоднородной системы по формуле

$$x^h = A_o^{-1} b = T \left(\begin{array}{c|c} E_r & O \\ \hline O & O \end{array} \right) S b$$

4. Составить фундаментальную матрицу

$$\Phi = T \left(\begin{array}{c} O \\ \hline E_{n-r} \end{array} \right)$$

5. Записать общее решение системы в виде

$$x = x^h + \Phi c,$$

где $c = (C_1 \dots C_{n-r})^T$ – столбец произвольных постоянных.

10. Псевдорешения системы линейных уравнений

Система m линейных алгебраических уравнений с n неизвестными $Ax = b$ может иметь единственное решение, бесконечно много решений или вообще не иметь решений. Нужно изменить понятие решения так, чтобы любая система линейных уравнений имела бы единственное в некотором смысле «решение».

Поставим каждому столбцу в соответствие неотрицательное действительное число, а именно норму (модуль)

$$|x| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}.$$

Псевдорешением системы линейных уравнений называется наименьший по норме столбец \tilde{x} среди всех столбцов, минимизирующих величину $|Ax - b|$.

ЗАМЕЧАНИЕ: *любая* система имеет единственное псевдорешение [1, стр. 209]

$$\tilde{x} = A^{\sim 1}b,$$

где $A^{\sim 1}$ – псевдообратная матрица для матрицы системы.

Понятие псевдорешения позволяет обойти не только факт неединственности, но и факт несуществования решений.

Если система несовместна, то псевдорешение \tilde{x} обеспечивает наименьшую величину погрешности $\varepsilon(x) = |Ax - b|$.

Если система совместна, то псевдорешение \tilde{x} является ее решением, т.е. $\varepsilon(\tilde{x}) = 0$, причем наименьшим по норме.

Алгоритм нахождения псевдорешения неоднородной системы:

1. Найти псевдообратную матрицу $A^{\sim 1}$.
2. Найти псевдорешение $\tilde{x} = A^{\sim 1}b$.

ЗАМЕЧАНИЕ: *полуобратная* матрица определена неоднозначно и потому задает не конкретное решение, а *множество решений* системы. *Псевдорешение*, полученное с помощью псевдообратной матрицы, всегда вычисляется в *конкретное решение*.

11. Свойства решений однородной системы

Общее решение однородной системы $Ax = o$ имеет вид [1, стр. 194]

$$\begin{cases} x_1 = -a'_{1,r+1}x_{r+1} - \dots - a'_{1,n}x_n, \\ \dots \\ x_r = -a'_{r,r+1}x_{r+1} - \dots - a'_{r,n}x_n. \end{cases}$$

Некоторые свойства:

- Если столбцы $\varphi_1, \varphi_2, \dots, \varphi_k$ – решения однородной системы уравнений, то любая их линейная комбинация $\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \dots + \alpha_k \varphi_k$ также является решением однородной системы,
- Если ранг матрицы однородной системы равен r , то система имеет $(n - r)$ *линейно независимых решений*.

Любая совокупность $(n - r)$ линейно независимых решений $\varphi_1, \varphi_2, \dots, \varphi_{n-r}$ однородной системы называется *фундаментальной системой решений*.

Теорема об общем решении однородной системы. Если $\varphi_1, \varphi_2, \dots, \varphi_{n-r}$ – фундаментальная система решений однородной системы уравнений, то столбец

$$x = C_1\varphi_1 + C_2\varphi_2 + \dots + C_{n-r}\varphi_{n-r} \quad (1)$$

при любых значениях произвольных постоянных C_1, C_2, \dots, C_{n-r} также является решением системы $Ax = 0$, и, наоборот, для каждого решения x этой системы найдутся такие значения произвольных постоянных C_1, C_2, \dots, C_{n-r} , при которых это решение x удовлетворяет равенству (1).

12. Функциональные матрицы скалярного аргумента

Функциональной матрицей скалярного аргумента t называется матрица, элементы которой являются функциями независимой переменной t

$$A(t) = [a_{ij}(t)]_{i,j=1}^{m,n}$$

Производная функциональной матрицы

$$\frac{dA(t)}{dt} = \left[\frac{da_{ij}(t)}{dt} \right]_{i,j=1}^{m,n}.$$

Производная обратной матрицы (если она существует)

$$\frac{dA^{-1}(t)}{dt} = -A^{-1}(t) \frac{dA(t)}{dt} A^{-1}(t).$$

Производная определителя квадратной матрицы $A(t)$ n -ого порядка

$$\frac{d}{dt} \det A(t) = \sum_{i=1}^n \sum_{j=1}^n A_{ij}(t) \frac{da_{ij}(t)}{dt} = \operatorname{tr} \left[A^+(t) \frac{dA(t)}{dt} \right],$$

где $A_{ij}(t)$ – алгебраическое дополнение элемента $a_{ij}(t)$ матрицы $A(t)$; $A^+(t)$ – присоединенная матрица.

13. Производные скалярной функции по векторному аргументу

Рассмотрим скалярную (числовую) функцию нескольких переменных $f(x_1, x_2, \dots, x_n)$. Упорядоченный набор переменных x_1, x_2, \dots, x_n будем называть *векторным аргументом* этой функции.

Первый дифференциал функции $f(x) = f(x_1, x_2, \dots, x_n)$ имеет вид

$$df(x) = \frac{\partial f(x)}{\partial x_1} dx_1 + \frac{\partial f(x)}{\partial x_2} dx_2 + \dots + \frac{\partial f(x)}{\partial x_n} dx_n.$$

Сумму в правой части можно представить как произведение строки $\frac{df(x)}{dx} = \left(\frac{\partial f(x)}{\partial x_1} \dots \frac{\partial f(x)}{\partial x_n} \right)$ на столбец $dx = (dx_1 \dots dx_n)^T$, либо как произведение строки dx^T на столбец $dx = \frac{df(x)}{dx^T} = \left(\frac{df(x)}{dx} \right)^T$.

Так как первый дифференциал $\underset{1 \times 1}{df(x)}$ – это одноэлементная матрица (а одноэлементная матрица совпадает со своей транспонированной), то $\underset{1 \times 1}{df(x)} = \underset{1 \times 1}{(df(x))^T}$

$$\underset{1 \times 1}{df(x)} = \underset{1 \times n}{\frac{df(x)}{dx}} \underset{n \times 1}{dx} = \left(\underset{1 \times n}{\frac{df(x)}{dx}} dx \right)^T = dx^T \left(\underset{1 \times n}{\frac{df(x)}{dx}} \right)^T = \underset{1 \times n}{dx}^T \underset{n \times 1}{\frac{df(x)}{dx^T}}.$$

Второй дифференциал функции имеет вид

$$d^2 f(x) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x)}{\partial x_i \partial x_j} dx_i dx_j.$$

Обозначим через $\frac{d^2 f(x)}{dx^T dx} = \left[\frac{\partial^2 f(x)}{\partial x_i^2 \partial x_j^2} \right]_{i,j=1}^n$ квадратную матрицу частных производных второго порядка (*матрицу Гессе*). Определитель матрицы Гессе называется *гессианом*.

Тогда можно переписать

$$\underset{1 \times 1}{d^2 f(x)} = \underset{1 \times n}{dx}^T \underset{n \times n}{\frac{d^2 f(x)}{dx^T dx}} \underset{n \times 1}{dx}.$$

Для скалярной функции скалярного аргумента второй дифференциал будет иметь вид

$$d^2 f(x) = \frac{d^2 f(x)}{dx^2} dx^2.$$

Для записи производных можно использовать символические векторы (столбцы или строки)

$$\nabla = \frac{d}{dx} = \left(\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_n} \right), \quad \nabla^T = \frac{d}{dx^T} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}.$$

При этом дифференцирование функции формально записывается как умножение функции на символический вектор производных. Например, градиент функции есть произведение вектора ∇ на функцию $f(x)$

$$\begin{aligned} \nabla f(x_1, \dots, x_n) &= \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n} \right), \\ \underset{n \times 1}{\nabla}^T \underset{1 \times n}{\nabla} &= \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \cdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \left(\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_n} \right) = \left[\frac{\partial^2}{\partial x_i \partial x_j} \right]_{i,j=1}^n. \end{aligned}$$

Найти первую и вторую производные сложной функции $g(t) = f(x_1(t), \dots, x_n(t))$, применяя матричные обозначения.

Находим производные функции, заменяя суммирование операциями умножения соответствующих матриц. Первая производная

$$\frac{dg(t)}{dt} = \frac{d}{dt} \left(f(x_1(t), \dots, x_n(t)) \right) = \sum_{i=1}^n \frac{\partial f(x(t))}{\partial x_i} \cdot \frac{dx_i(t)}{dt} = \frac{df(x(t))}{dx} \cdot \frac{dx(t)}{dt}$$

В случае скалярной функции скалярного аргумента первая производная от функции $g(x)$ будет выглядеть так же.

Вторая производная скалярной функции векторного аргумента

$$\begin{aligned} \frac{d^2 g(t)}{dt^2} &= \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f(x(t))}{\partial x_j \partial x_i} \cdot \frac{dx_i(t)}{dt} \cdot \frac{dx_j(t)}{dt} + \sum_{i=1}^n \frac{\partial f(x(t))}{\partial x_i} \cdot \frac{dx_i^2(t)}{dt^2} = \dots \\ &\dots = \left(\frac{dx(t)}{dt} \right)^T \frac{d^2 f(x(t))}{dx^T dx} \cdot \frac{dx(t)}{dt} + \frac{df(x(t))}{dx} \cdot \frac{d^2 x(t)}{dt^2}. \end{aligned}$$

В случае скалярной функции скалярного аргумента вторая производная будет выглядеть так

$$\frac{d^2 g(t)}{dt^2} = \frac{d^2 f(x(t))}{dx^2} \left(\frac{dx(t)}{dt} \right)^2 + \frac{df(x(t))}{dx} \cdot \frac{d^2 x(t)}{dt^2}.$$

Выражения для первой производной совпадают, а для второй производной – отличаются незначительно, причем полное совпадение будет, если учесть, что $x^T = x$ для скалярной величины x .

14. Производные от векторной функции векторного аргумента

Пусть задан столбец

$$f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}$$

функций нескольких переменных (говорят, что задана *вектор-функция векторного аргумента*).

Первый дифференциал вектор-функции имеет вид

$$df(x) = \begin{pmatrix} df_1(x_1, \dots, x_n) \\ \vdots \\ df_m(x_1, \dots, x_n) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \frac{\partial f_1(x_1, \dots, x_n)}{\partial x_j} dx_j \\ \vdots \\ \sum_{j=1}^n \frac{\partial f_m(x_1, \dots, x_n)}{\partial x_j} dx_j \end{pmatrix} = \sum_{j=1}^n \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_j} \\ \vdots \\ \frac{\partial f_m(x)}{\partial x_j} \end{pmatrix} dx_j$$

Обозначим через

$$\frac{df(x)}{dx} = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \dots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix} = \left[\frac{\partial f_i(x)}{\partial x_j} \right]_{i,j=1}^{m,n}$$

матрицу частных производных первого порядка заданных функций (*матрицу Якоби*).

Тогда выражение для первого дифференциала можно записать в виде $df(x) = \frac{df(x)}{dx} dx$, т.е. $\frac{df(x)}{dx}$ – производная вектор-функции векторного аргумента.

Как и в случае с аргументом x , упорядоченный набор функций можно считать не матрицей-столбцом, а матрицей-строкой $(f(x))^T$. Этот случай сводится к предыдущему, учитывая, что операции дифференцирования и транспонирования можно выполнять в любом порядке, так как $d(f^T) = (df)^T$. Тогда из равенства $df = \frac{df}{dx} dx$ получаем $df^T = (dx)^T \left(\frac{df}{dx} \right)^T = (dx)^T \frac{df^T}{dx^T}$, где

$$\left(\frac{df(x)}{dx} \right)^T = \frac{df^T}{dx^T} = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1(x)}{\partial x_n} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix}$$

– транспонированная матрица Якоби вектор-функции векторного аргумента.

15. Правила дифференцирования по векторному аргументу

Векторный аргумент x , его приращение dx считаем вектор-столбцами размеров $n \times 1$.

Первый дифференциал скалярной функции векторного аргумента $f(x_1, \dots, x_n)$ (одноэлементная матрица) имеет вид

$$df = \frac{df}{dx} dx = dx^T \frac{df}{dx^T},$$

где $\frac{df}{dx} = \left(\frac{df}{dx_1} \dots \frac{df}{dx_n} \right)$ – градиент функции, а $\left(\frac{df}{dx} \right)^T = \frac{df}{dx^T}$, так как функция скалярная.

Второй дифференциал скалярной функции векторного аргумента $f(x_1, \dots, x_n)$

$$d^2 f = dx^T \frac{d^2 f}{dx^T dx} dx,$$

где $\frac{d^2 f}{dx^T dx} = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right]_{i,j=1}^n$ – матрица Гессе.

Первый дифференциал вектор-функции векторного аргумента (вектора-столбца) $f(x)$ имеет вид

$$df = \frac{df(x)}{dx} dx,$$

где $\frac{df(x)}{dx}$ – матрица Якоби.

Первый дифференциал вектора-строки

$$(df)^T = d(f^T) = dx^T \frac{df^T}{dx^T}.$$

В частном случае, когда $f(x_1, \dots, x_n) = (x_1, \dots, x_n)$, получаем

$$\frac{dx}{dx} = E, \quad \frac{dx^T}{dx^T} = E,$$

где E – единичная матрица n -ого порядка.

Числовую матрицу C соответствующих размеров можно выносить за знак производной

$$\frac{d(Cf)}{dx} = C \frac{df}{dx}, \quad \frac{d(f^T C)}{dx^T} = \frac{df^T}{dx^T} C$$

Производные суммы, разности и произведения вектор-функций векторного аргумента $u(x)$ и $v(x)$ одинаковых размеров $m \times 1$

$$\begin{aligned} \frac{d(u+v)}{dx} &= \frac{du}{dx} + \frac{dv}{dx}, & \frac{d(u-v)}{dx} &= \frac{du}{dx} - \frac{dv}{dx}, \\ \frac{d(u^T v)}{dx} &= u^T \frac{dv}{dx} + v^T \frac{du}{dx}, & \frac{d(u^T v)}{dx^T} &= \frac{du^T}{dx^T} v + \frac{dv^T}{dx^T} u. \end{aligned}$$

Производная сложной функции $z(y(x))$, где $z = z(y) = \begin{pmatrix} z_1(y) \\ \vdots \\ z_k(y) \end{pmatrix}$ и $y = y(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_m(x) \end{pmatrix}$,

вычисляется по формуле $\frac{dz(y(x))}{dx} = \frac{dz(y(x))}{dy} \frac{dy(x)}{dx}$ или, опуская аргументы, $\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}$.
 $k \times m \quad m \times n$

След матрицы Якоби (при $m = n$) определяет *дивергенцию*

$$\operatorname{div} f = \operatorname{tr} \frac{df}{dx} = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i},$$

где $f(x)$ – векторная функция векторного аргумента.

16. Производные матричной функции по векторному аргументу

Рассмотрим функциональную матрицу $A(x)$, элементами которой служат скалярные функции $a_{ij}(x)$ векторного аргумента x . То есть такая матрица представляет собой трехмерную сущность, в которой на пересечении строки и столбца стоит скалярная функция, имеющая «глубину» в виде вектора аргументов.

Первый дифференциал этой функции

$$dA(x) = \sum_{i=1}^n \frac{\partial A(x)}{\partial x_i} dx_i,$$

где $\frac{\partial A(x)}{\partial x_i}$ – частная производная матрицы по одной переменной.

Совокупность частных производных (градиент функциональной матрицы) представляет собой объект, элементы которого $\frac{\partial a_{ij}(x)}{\partial x_k}$ нумеруются тремя индексами: номер строки, номер столбца и номер переменной дифференцирования. Поэтому заменить операцию суммирования в правой части формулы операций умножения матриц в данном случае не представляется возможным. Необходимо вводить тензоры и операции над ними.

Элементы матрицы $A = (a_j^i)$ обозначаются a_j^i , где i – номер строки, а j – номер столбца. В частности, $x = (x^i)$ – столбец, а $y = (y_j)$ – строка.

Частную производную функции $F(x)$ (скалярной, векторной или матричной), то по ним производится суммирование (хотя знак суммы не указывается). Например, если $A = (a_j^i)$ – матрица

размеров $m \times n$, $x = (x^j)$ – столбец размеров $n \times 1$, $y = (y_i)$ – строка размеров $1 \times m$, то

$$a_j^i x^j = \sum_{j=1}^n a_j^i x^j, \quad a_j^i y_i = \sum_{i=1}^m a_j^i y_i, \quad a_j^i x^j y_i = \sum_{i=1}^m \sum_{j=1}^n a_j^i x^j y_i,$$

т.е. $a_j^i x^j$ – i -ый элемент столбца Ax ; $a_j^i y_i$ – j -ый элемент строки yA ; $a_j^i x^j y_i$ – число yAx .

Применяя эти соглашения, запишем дифференциалы

$$\begin{aligned} df &= f_{(i)} dx^i, & d^2 f &= f_{(i)(j)} dx^i dx^j, \\ df^i &= f_{(j)}^i dx^j, \\ df_j^i &= f_{j(k)}^i dx^k, \end{aligned}$$

где $f_{j(k)}^i = \frac{\partial f_j^i}{\partial x^k}$ – частная производная первого порядка элемента f_j^i функциональной матрицы F по переменной x^k .

17. Линейные и квадратичные формы

Многочлен первой степени от n переменных x_1, \dots, x_n называется выражением вида

$$p_1(x) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n + c_0,$$

где c_0, c_1, \dots, c_n – коэффициенты многочлена (предполагается, что среди коэффициентов есть отличные от нуля); коэффициент c_0 называется свободным членом. Многочлен первой степени называется однородным, если $p_1(\lambda x) = \lambda p_1(x)$ для любого числа λ (это возможно только когда $c_0 = 0$).

Линейной формой переменных x_1, \dots, x_n называется однородный многочлен первой степени

$$g(x) = \sum_{i=1}^n c_i x_i,$$

где $\{c_i\}_{i=1}^n$ – коэффициенты линейной формы.

Составляя из коэффициентов строку $c = (c_1 \dots c_n)$, а из переменных – столбец $x = (x_1 \dots x_n)^T$, линейную форму можно записать в виде

$$g(x) = cx.$$

Многочлен второй степени от n переменных x_1, \dots, x_n называется выражением

$$p_2(x) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c_0,$$

где числа a_{ij}, b_i, c_0 – коэффициенты многочлена: a_{ij} – квадратичные коэффициенты; b_i – коэффициенты линейных членов; c_0 – свободный член.

Многочлен второй степени называется однородным, если $p_2(\lambda x) = \lambda^2 p_2(x)$ (это возможно только когда $b_1 = b_2 = \dots = b_n = 0, c_0 = 0$).

Квадратичной формой переменных x_1, \dots, x_n называется однородный многочлен второй степени

$$q(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j, \quad a_{ij} = a_{ji}.$$

Симметрическая матрица $A = (a_{ij})$, составленная из коэффициентов квадратичной формы, называется матрицей квадратичной формы.

Квадратичная форма называется вырожденной, если ее матрица вырожденная ($\text{rg } A < n$), в противном случае, когда матрица невырожденная ($\text{rg } A = n$), квадратичная форма называется невырожденной.

Составляя из переменных столбец $x = (x_1 \dots x_n)^T$, квадратичную форму можно записать в виде

$$q(x) = x^T A x.$$

Список литературы

1. Бортакровский А.С. Линейная алгебра в примерах и задачах. – М.: Высш. шк., 2005. – 591 с.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1972. – 368 с.
3. Бахвалов Н. С. Численные методы. – М.: Лаборатория Базовых Знаний, 2000. – 624 с.
4. Шевцов Г.С. Численные методы линейной алгебры. – М.: Финансы и статистика, 2012. – 480 с.
5. Лагутин М.Б. Наглядная математическая статистика. – М.: БИНОМ, 2009. – 472 с.
6. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2012. – 816 с.