Дата последней сборки документа:
22 июля 2021 г. в 21:59

Сборник заметок по линейной алгебре и сопряженным вопросам

Подвойский А.О.

Содержание

1. 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

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

Тогда решение системы представляется столбцом

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

и удовлетворяте равенству

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

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

2. 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 линейный независимых уравнений с n неизвестными отличен от нуля ($\det A \neq 0$), то система имеет *единственное* решение, которое находится по формулам

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

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 $\Delta = \Delta_1 = \Delta_2 = \ldots, \Delta_n = 0$, то возможны два случая: либо система несовместна (не имеет ни одного решения), либо система имеет бесконечно много решений [1, стр. 188].

3. 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*

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

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 $\operatorname{rg} A = \operatorname{rg}(A \mid b)$.

Если $\operatorname{rg} A \neq \operatorname{rg}(A \mid b)$, 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, $\operatorname{rg}(A \mid b) = \operatorname{rg} A$.

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

4. Общее решение системы линейных алгебраических уравнений

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

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

$$\begin{cases} x_1 = b_1' - a_{1,r+1}' x_{r+1} - \ldots - a_{1,n}' x_n, \\ \ldots \\ x_r = b_r' - a_{r,r+1}' x_{r+1} - \ldots - a_{r,n}' x_n, \end{cases}$$

где $x_1, x_2, \ldots, x_r$ – базисные переменные; $x_{r+1}, x_{r+2}, \ldots, x_n$ – 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* <u>неоднородной</u> системы и *общего решения* соответствующей <u>однородной</u> системы [1, стр. 200]

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

5. Решение систем уравнений с помощью полуобратных матриц

Требуется решить систему линейных уравнений

$$Ax = b,$$

где A – <u>произвольная</u> 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 $A^{\neg 1}$ – матрица, полуобратная к матрице системы A. Используя определение полуобратной матрицы, неоднородную систему $Ax = b$ можно переписать так

$$AA^{\neg 1}Ax = b.$$

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

$$AA^{\neg 1}Ax = b, \quad \to \quad AA^{\neg 1}b = b.$$

Тогда

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

Это необходимое и достаточное условие совместности системы.

Решением системы будет $x = A^{\neg 1}b$. 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

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

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

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]

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

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

Алгоритм применения полуобратной матрицы:

1. Привести матрицу A системы $Ax = b$ к простейшему виду: $\Lambda = SAT$. При этом находятся элементраные преобразующие матрицы S и T, а также ранг $r = \operatorname{rg} A \geqslant 1$.
2. Проверить условие совместности системы $\Psi b = o$. При $r = m$ система совместна. Если $r < m$, то составить матрицу $\Psi = (O \,|\, E_{m-r})\, S$ 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
$$x^{\text{н}} = A_o^{\neg 1}\, b = T \left(\begin{array}{c|c} E_r & O \\ \hline O & O \end{array}\right) S\, b$$
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

$$x = x^{\text{н}} + \Phi\, c,$$

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

6. 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 соответсвие неотрицательное действительное число, а именно норму (модуль)

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

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.*

7. Свойства решений однородной системы

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

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

Некоторые свойства:

- Если столбцы $\varphi_1, \varphi_2, \ldots, \varphi_k$ – решения однородной системы уравнений, то любая их линейная комбинация $\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \ldots + \alpha_k \varphi_k$ 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 $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ однородной системы называется *фундаментальной системой решений.*

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

$$x = C_1\varphi_1 + C_2\varphi_2 + \ldots + C_{n-r}\varphi_{n-r} \tag{1}$$

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

8. Функциональные матрицы скалярного аргумента

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

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

Производная функциональной матрицы

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

Производная обратной матрицы (если она существует)

$$\underset{m\times n}{\frac{dA^{-1}(t)}{dt}} = -A^{-1}(t)\,\frac{dA(t)}{dt}\,A^{-1}(t).$$

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.

9. Производные скалярной функции по векторному аргументу

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

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

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

Сумму в правой части можно представить как произведение строки $\frac{df(x)}{dx} = \left(\frac{\partial f(x)}{\partial x_1} \ldots \frac{\partial f(x)}{\partial x_n}\right)$ на столбец $dx = (dx_1 \ldots dx_n)^T$, 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 одноэлементая матрица совпадает со своей транспонированной), то $\underset{1\times 1}{df(x)} = \big(\underset{1\times 1}{df(x)}\big)^T$

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

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)

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

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

$$\nabla f(x_1,\ldots,x_n) = \Big(\frac{\partial f}{\partial x_1}\ldots\frac{\partial f}{\partial x_n}\Big),$$

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

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

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

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

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

Вторая производная скалярной функции векторного аргумента

$$\frac{d^2g(t)}{dt^2} = \sum_{j=1}^n\sum_{i=1}^n \frac{\partial^2 f(x(t))}{\partial x_j \partial x_i}\cdot\frac{dx_i(t)}{dt}\cdot\frac{dx_j(t)}{dt} + \sum_{i=1}^n \frac{\partial f(x(t))}{\partial x_i}\cdot\frac{dx_i^2}{dt^2} = \ldots$$
$$\ldots = \Big(\frac{dx(t)}{dt}\Big)^T \frac{d^2 f(x(t))}{dx^T dx}\cdot\frac{dx(t)}{dt} + \frac{df(x(t))}{dx}\cdot\frac{d^2x(t)}{dt^2}.$$

В случае скалярной функции скалярного аргумента вторая производная будет выглядеть так

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

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.

Список литературы

1. *Бортаковский 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. *Лагутин М.Б.* Наглядная математическая статистика. – М.: БИНОМ, 2009. – 472 с.
4. *Кобзарь А.И.* Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2012. – 816 с.