矩阵半张量积的基本原理与适用领域*

程代展, 齐洪胜

中国科学院数学与系统科学研究院系统科学研究所系统控制重点实验室, 北京 100190

E-mail: dcheng@iss.ac.cn, qihongsh@amss.ac.cn

摘要 作为一种新的矩阵乘法, 矩阵半张量积正得到国内外学者越来越多的重视和参与, 从而使之应用于越来越多的研究课题中. 本文希望分析矩阵半张量积的基本原理, 从其合理性说明它产生的必然性和存在的意义. 同时, 与已有的综述不同, 本文不准备具体介绍它在那些问题中得到那些应用, 而是从原理出发, 说明它可能在那些类型的相关科学问题中得到应用. 这使我们能够更主动地去开发它可能的潜在应用.

关键词 矩阵半张量积, 布尔网络, 泛代数, 多元多项式, 非线性矩阵方法状态空间与正则子空间

1 引言

"长期以来, 线性代数与矩阵理论一直是许多数学分支的基本工具, 同时, 它们自身也具有丰富的研究课题." (Linear algebra and matrix theory have long been fundamental tools in mathematical disciplines as well as fertile fields for research in their own right.) [6] 相信每一个从事数学甚至其他自然科学的学者都不会怀疑矩阵的重要性. 它和微积分可以算是数学的两块基石, 大致可以说, 整个近代数学的大厦是建立在这两大基石之上的. 虽然矩阵的观念十分直观, 它远比微积分早出现, 例如我国出现于公元前一世纪的《九章算术》, 就已经把线性方程组的系数排成方阵 (矩阵), 进而用消元法 (即高斯消去法) 求解 [7]. 但与微积分不同, 矩阵理论还在不断发展, 正如 [6] 指出的: 矩阵自身具有丰富的研究课题. 至今还有许多关于矩阵/线性代数的国际期刊, 每年都有成百上千的新结果.

为叙述方便计, 先规定一些记号:

(i) $\mathcal{M}_{m\times n}$: $m\times n$ 实矩阵集合.

(ii)
$$\mathcal{D}_k=\left\{0,\frac{1}{k-1},\frac{2}{k-1},\cdots,1\right\},\quad \mathcal{D}_2=\mathcal{D}=\{0,1\}.$$

(iii) δ_n^k: 单位阵 I_n 的第 k 列.

(iv) $\Delta_n:=\{\delta_n^1,\cdots,\delta_n^n\},\quad \Delta:=\Delta_2$.

(v) $\mathrm{Col}_i(A)$ ($\mathrm{Row}_i(A)$): 矩阵 A 的第 i 行 (列). 矩阵 A 的行 (列) 集合记作 $\mathrm{Col}(A)$ ($\mathrm{Row}(A)$).

(vi) 矩阵 $L\in M_{n\times m}$ 如果它的列满足 $\mathrm{Col}(M)\subset\Delta_n$, 则称其为逻辑矩阵. $n\times m$ 维逻辑矩阵集合记作 $\mathcal{L}_{n\times m}$.

*国家自然科学基金 (61074114, 60821091, 61104065) 资助课题.

(vii) 设 $L \in \mathcal{L}_{n\times m}$. 则

$$L = [\delta_n^{i_1}, \delta_n^{i_2}, \cdots, \delta_n^{i_m}],$$

简记为

$$L = \delta_n[i_1, i_2, \cdots, i_m].$$

(viii) 记 $\mathbf{1}_n$ 为一维列向量, 其元素均为 1.

(ix) $W_{[n,m]} \in \mathcal{M}_{mn\times mn}$ 称为换位矩阵, 如果

$$W_{[m,n]}(X \otimes Y) = Y \otimes X, \quad X \in \mathbb{R}^m,\ Y \in \mathbb{R}^n. \tag{1}$$

显然, 矩阵运算还不能说已经尽善尽美, 没有缺点. 如果和数的乘法相比, 普通矩阵乘法的一个最大缺点是: 只有前阵列数与后阵行数相等, 矩阵乘法才有效. 因此, 即使从数学本身出发, 一个很自然的问题是: 能不能把矩阵普通乘法推广到任意两个矩阵? 矩阵半张量积正是这样一个推广. 它可以一般地定义如下 [3]: 设 $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{p\times q}$. 记 $t = \mathrm{lcm}\{n, p\}$, 即 t 为 n 和 p 的最小公倍数, 那么 A 和 B 的半张量积, 记作 $A \ltimes B$, 定义为

$$A \ltimes B := \left(A \otimes I_{t/n}\right)\left(B \otimes I_{t/p}\right). \tag{2}$$

我们将 (2) 称为矩阵左半张量积, 通常说矩阵半张量积均指左半张量积. 容易看出, (2) 是普通矩阵乘法的推广, 因为当 $n = p$ 时, 它就是普通矩阵乘法. 其实, 这种推广可以有很多. 例如, 右半张量积 [3]:

$$A \rtimes B := \left(I_{t/n} \otimes A\right)\left(I_{t/p} \otimes B\right). \tag{3}$$

当然, 也可以定义混合半张量积

$$A \bowtie B := \left(I_{t/n} \otimes A\right)\left(B \otimes I_{t/p}\right); \tag{4}$$

或

$$A \bowtie B := \left(A \otimes I_{t/n}\right)\left(I_{t/p} \otimes B\right). \tag{5}$$

还有一个有几何意义的推广:

$$A \odot B := \left(\mathbf{1}_{t/n} \otimes A\right)\left(\mathbf{1}_{t/p} \otimes B\right). \tag{6}$$

此外, [10] 还将左半张量积推广为泛张量积, 等等. 听一位老院士说: "华罗庚先生曾提出过一种推广, 即如果 $n < p$, 给 A 补上 $p - n$ 个零元素的列, 如果 $n > p$, 给 B 补上 $n - p$ 个零元素的行." 这大概是最简单的推广了. 其实, 学过线性代数的大学生, 只要有兴趣, 都不难造出自己的推广. 这位院士还说: "华先生说过, 把矩阵乘法推广到任意两个矩阵没有意义." 我们理解, 华先生指的是上述或他心中想到的一些推广. 从正面理解华先生的意思, 应当是: 不要随便去做没有意义的推广. 既然华先生提到了推广, 他大概也探讨过可能的合理或有意义的推广.

什么是合理且有意义的推广, 我们以为大致有两条:

(i) 它必须定义简明、计算方便, 同时, 不能破坏矩阵乘法原有的基本性质;

(ii) 它必须有明确的物理意义, 从而它能够成为解决相应实际问题的强有力工具.

否则, 推广将只能是一个简单的数学游戏. 本文的第一个目的, 就是要说明, 矩阵的半张量积 (以下指定义 (2)) 满足上述两条.

2 矩阵半张量积的基本原理

设 $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{p\times q}$. 如果 $n = p$, 我们称 A 和 B 满足等维数条件; 如果 $n = ps$ (记作 $A \succ_s B$) 或 $ns = p$ (记作 $A \prec_s B$), 我们称 A 和 B 满足倍维数条件; 如果倍维数条件也不满足, 则称它们为两个任意矩阵. 实际上, 现在最有效的应用均属倍维数情形. 在这种情况下, 我们有如下等价定义:

定义 1 *(i)* 设 $X \in \mathbb{R}^{mn}$ 为一行向量, $Y \in \mathbb{R}^m$ 为一列向量. 将 X 等分成 m 段, 即 $X = (X^1\ X^2\ \cdots\ X^m)$, 这里 $X^i \in \mathbb{R}^n$, $i = 1, \cdots, m$. 那么, X 和 Y 的半张量积, 记作 $X \ltimes Y$, 定义为

$$X \ltimes Y := \sum_{i=1}^{m} X^i y_i \in \mathbb{R}^n. \tag{7}$$

(ii) 设 $X \in \mathbb{R}^m$ 为一行向量, $Y \in \mathbb{R}^{mn}$ 为一列向量. 那么, X 和 Y 的半张量积定义为

$$X \ltimes Y := \left(Y^T \ltimes X^T\right)^T \in \mathbb{R}^n. \tag{8}$$

(iii) 设矩阵 $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{p\times q}$ 满足倍维数条件, 那么, 利用 (7) 及 (8), 定义

$$A \ltimes B = \begin{bmatrix} \mathrm{Row}_1(A) \ltimes \mathrm{Col}_1(B) & \mathrm{Row}_1(A) \ltimes \mathrm{Col}_2(B) & \cdots & \mathrm{Row}_1(A) \ltimes \mathrm{Col}_q(B) \\ \mathrm{Row}_2(A) \ltimes \mathrm{Col}_1(B) & \mathrm{Row}_2(A) \ltimes \mathrm{Col}_2(B) & \cdots & \mathrm{Row}_2(A) \ltimes \mathrm{Col}_q(B) \\ \vdots & & & \\ \mathrm{Row}_m(A) \ltimes \mathrm{Col}_1(B) & \mathrm{Row}_m(A) \ltimes \mathrm{Col}_2(B) & \cdots & \mathrm{Row}_m(A) \ltimes \mathrm{Col}_q(B) \end{bmatrix}. \tag{9}$$

这个定义给出的物理概念是很清楚的.

实际上, 经典矩阵理论能方便处理线性或二次函数. 但对高次函数它却很难应付. 现在看它是如何处理二次函数, 或者说双线性形式. 设 $X \in U$ 且 $X \in V$, U 是一个 m 维向空间, 其基底为 $\{u_1, \cdots, u_m\}$, V 是一个 n 维向空间, 其基底为 $\{v_1, \cdots, v_n\}$. $f : U \times V \to \mathbb{R}$ 是一个双线性函数. 如果

$$f(u_i, v_j) = a_{i,j}, \quad i = 1, \cdots, m;\ j = 1, \cdots, n,$$

则我们可将 $\{a_{i,j}\}$ 排成一个矩阵

$$A = (a_{i,j}) \in \mathcal{M}_{m\times n}.$$

记 $X = \sum_{i=1}^{m} x_i u_i \sim (x_1, \cdots, x_m)^T$, $Y = \sum_{i=1}^{n} y_i v_i \sim (y_1, \cdots, y_n)^T$. 则在向量形式下, 我们有矩阵表示式

$$f(X, Y) = X^T A Y. \tag{10}$$

其实, 我们也可以换一种方法表示这个双线性映射. 将 $\{a_{i,j}\}$ 排成一个行, 记作

$$M_f = (a_{11},\ a_{12},\ \cdots,\ a_{1n},\ \cdots,\ a_{m1},\ a_{m2},\ \cdots,\ a_{mn}).$$

那么, 利用定义 1 不难验证

$$f(X, Y) = M_f \ltimes X \ltimes Y. \tag{11}$$

实际上, (10) 与 (11) 的目的都是一样的, 它要通过运算规则让变量 $\{x_i, y_j\}$ 找到它们对应的系数 a_{ij}.

现在考虑三线性映射. 设 U, V 如前, W 是一个 s 维向空间, 其基底为 $\{w_1, \cdots, w_s\}$. $g: U \times V \times W \to \mathbb{R}$ 是一个三线性函数. 如果

$$g(u_i, v_j, w_k) = b_{i,j,k}, \quad i = 1, \cdots, m;\ j = 1, \cdots, n;\ k = 1, \cdots, s,$$

那么 $\{a_{i,j,k}\}$ 是一个 3 维数组. 对于三维数组, 能否像双线性函数 f 那样, 用矩阵的方法计算 g 呢? 实际上, 在许多科学计算中, 特别是统计问题中, 经常需要处理 3 维数组, 于是, 有人提出立方积的概念. 即把 3 维数组排成一个立方体 (如图 1), 然后定义相应的一套运算 [1,9]. 虽然立方积确实得到了一些应用, 但它需要一套复杂的运算规则. 况且, 它不可能推广到更高维的数组.

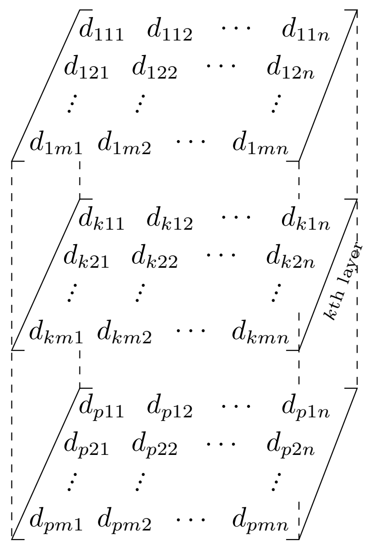

图 1: 立体阵

那么, 半张量积的方法是否可行呢? 将 $\{b_{i,j,k}\}$ 排成一个行, 记作

$$M_g = (b_{111},\ b_{112},\ \cdots,\ b_{11s},\ \cdots,\ b_{mn1},\ b_{mn2},\ \cdots,\ b_{mns}).$$

不难检验

$$g(X, Y, Z) = M_g \ltimes X \ltimes Y \ltimes Z. \tag{12}$$

实际上, 这种方法可以直接推广到多线性映射的情况, 这就是半张量积的本质所在. 在 (12) 中, 可以首先计算 $M_g \ltimes X$, 这时, 依半张量积的定义, M_g 中的 b_{i**} 会自动与 x_i 相乘; 再考虑 $M_g \ltimes X \ltimes Y$, 这时, M_g 中的 b_{ij*} 会自动与 x_i 及 y_j 相乘. 最后 $M_g \ltimes X \ltimes Y \ltimes Z$ 会让 b_{ijk} 自动与 x_i, y_j 及 z_k 相乘. 一般地说, 半张量积可自动寻找多线性映射中每个变量所对应的系数集, 因此, 它可以方便地用于多线性映射.

矩阵半张量积保持了矩阵普通乘法的所有重要性质, 因此, 在应用上无需区分普通积和半张量积, 可对任意两矩阵按普通积一样操作. 不仅如此, 矩阵半张量积有许多伪交换性, 它在一定程度上克服了矩阵乘法的不可交换性. 这使半张量积在应用上比普通乘法更方便. 这些, 使半张量积有了被广泛应用的可能.

但是, 矩阵半张量积的真正魅力是在于它前面指出的那种自动寻找变量组 $\{x_i, y_j, z_k, \cdots\}$ 与系数 $\mu_{ijk\cdots}$ 的对应关系上. 半张量积的所有应用都基于它的这种自动寻配能力上.

下面讨论半张量积可能的应用领域.

3 有限集间的映射

设有三个有限集 $U=\{u_1,\cdots,u_m\}$, $V=\{v_1,\cdots,v_n\}$, $W=\{w_1,\cdots,w_s\}$. 函数 $f:U\times V\to W$, 满足

$$f(u_i,v_j)=w_{k(i,j)},\quad i=1,\cdots,m;\ j=1,\cdots,n. \tag{13}$$

为了用矩阵表示 f, 我们用以下向量表示各元素:

$$\begin{array}{ll} u_i\sim\delta_m^i, & i=1,\cdots,m\\ v_i\sim\delta_n^i, & i=1,\cdots,n\\ w_i\sim\delta_s^i, & i=1,\cdots,s.\end{array}$$

则 f 可表示为

$$f(u,v)=M_f uv, \tag{14}$$

这里

$$M_f=\delta_s[k(1,1)\ k(1,2)\ \cdots\ k(1,n)\ \ \cdots\ \ k(m,1)\ k(m,2)\ \cdots\ k(m,n)]\in\mathcal{L}_{s\times mn},$$

称为 f 的结构矩阵. 表达式 (14) 的优势也在于它能够自动寻配. 例如, 固定 $u=u_0$, 则

$$f_0:=f|_{u=u_0}:V\to W.$$

它的结构矩阵为 $M_{f_0}=M_f u_0$. 如果固定 $v=v_0$, 则

$$f^0:=f|_{v=v_0}:U\to W.$$

因为

$$f(u,v)=M_f uv=M_f W_{[n,m]}vu.$$

则 f^0 的结构矩阵为 $M_{f^0}=M_f W_{[n,m]}v_0$.

推广到一般情况, 设 $V_i=\{v_1^i,v_2^i,\cdots,v_{n_i}^i\}$, $i=1,\cdots,k$, $W=\{w_1,w_2,\cdots,w_{n_0}\}$. $\pi:\prod_{i=1}^k V_i\to W$.

$$\pi(v_{t_1}^1,\cdots,v_{t_k}^k)=w_{s(t_1,\cdots,t_k)},\quad t_i=1,\cdots,n_i,\ i=1,\cdots,k. \tag{15}$$

那么

$$\pi(v_1,\cdots,v_k)=M_f v_1 v_2\cdots v_k, \tag{16}$$

这里

$$M_\pi=\delta_{n_0}\left[s(1,\cdots,1),s(1,\cdots,2),\cdots,s(1,\cdots,n_k),\cdots,s(n_1,n_2\cdots,1),\cdots,s(n_1,n_2\cdots,n_k)\right]. \tag{17}$$

现在如果 $v_s=v_s^0$, 定义

$$\pi_0=\pi|_{v_s=v_s^0}.$$

则 π_0 的结构矩阵为

$$M_{\pi_0}=M_\pi W_{[n_s,n^s]}v_s^0, \tag{18}$$

这时 $n^s = n_1 n_2 \cdots n_{s-1}$.

当考虑布尔网络时, 我们有

$$\begin{cases} x_1(t+1) = f_1(x_1(t), \cdots, x_n(t)) \\ \vdots \\ x_n(t+1) = f_n(x_1(t), \cdots, x_n(t)), \end{cases} \tag{19}$$

这里 $x_i \in \mathcal{D}$, $i = 1, \cdots, n$. 记 $x = \ltimes_{i=1}^n x_i$, 则存在唯一的 $L \in \mathcal{L}_{2^n \times 2^n}$, 使得 (19) 可表示为一个代数差分方程形式:

$$x(t+1) = Lx(t). \tag{20}$$

当考虑布尔控制网络时, 类似地, 我们有

$$\begin{array}{l} \begin{cases} x_1(t+1) = f_1(x_1(t), \cdots, x_n(t), u_1(t), \cdots, u_m(t)) \\ \vdots \\ x_n(t+1) = f_n(x_1(t), \cdots, x_n(t), u_1(t), \cdots, u_m(t)), \end{cases} \\ y_k(t) = h_k(x_1(t), \cdots, x_n(t)), \quad k = 1, \cdots, p. \end{array} \tag{21}$$

这里 $x_i, u_j, y_k \in \mathcal{D}$, $i = 1, \cdots, n$; $j = 1, \cdots, m$; $k = 1, \cdots, p$. 记 $x = \ltimes_{i=1}^n x_i$, $u = \ltimes_{j=1}^m u_j$, $y = \ltimes_{k=1}^p y_k$, 则存在唯一的 $L \in \mathcal{L}_{2^n \times 2^{n+m}}$, $H \in \mathcal{L}_{2^p \times 2^n}$, 使得 (21) 可表示为一个代数形式:

$$\begin{cases} x(t+1) = Lu(t)x(t) \\ y(t) = Hx(t). \end{cases} \tag{22}$$

这种代数形式成了整个基于半张量积的布尔网络控制理论基础 [4].

4 泛代数的半张量积方法

先从一个有限维代数说起. 设 V 为 $\mathbb{R}$ 上的一个 n 维空间, $\{v_1, \cdots, v_n\}$ 为它的一组基. 相对这组基, $X \in V$ 可表示为

$$X = \sum_{i=1}^n x_i v_i \sim (x_1, \cdots, x_n)^T.$$

如果在 V 上有一个乘法 $\odot : V \times V \to V$, 它是一个双线性映射, 那么 $\mathcal{V} := (V, \odot)$ 就构成一个代数. 如果

$$v_i \odot v_j = \sum_{k=1}^n \mu_{ij}^k v_k, \quad i, j = 1, \cdots, n.$$

那么, 矩阵

$$M_{\mathcal{V}} = \begin{bmatrix} \mu_{11}^1 & \cdots & \mu_{1n}^1 & \cdots & \mu_{n1}^1 & \cdots & \mu_{nn}^1 \\ \mu_{11}^2 & \cdots & \mu_{1n}^2 & \cdots & \mu_{n1}^2 & \cdots & \mu_{nn}^2 \\ \vdots & & & & & & \\ \mu_{11}^n & \cdots & \mu_{1n}^n & \cdots & \mu_{n1}^n & \cdots & \mu_{nn}^n \end{bmatrix}$$

称为代数 $\mathcal{V}$ 的结构矩阵.

一个代数的性质完全由它的结构矩阵所决定. 例如, 一个代数是李代数, 当且仅当, 它满足

(i) 反对称性

$$X \odot Y = -Y \odot X, \quad X, Y \in \mathcal{V}; \tag{23}$$

(ii) Jacobi 等式

$$(X \odot Y) \odot Z + (Y \odot Z) \odot X + (Z \odot X) \odot Y = 0, \quad X, Y, Z \in \mathcal{V}. \tag{24}$$

(23) 及 (24) 分别等价于结构矩阵满足以下的 (25) 及 (26) [3].

(i) 反对称性

$$M_{\mathcal{V}} \left[I_{n^2} + W_{[n,n]} \right] = 0; \tag{25}$$

(ii) Jacobi 等式

$$M_{\mathcal{V}}^2 \left[I_{n^3} + W_{[n,n^2]} + W_{[n^2,n]} \right] = 0. \tag{26}$$

简单地说, 向量空间上的一种运算 (如上述的 $\odot$) 称为一种代数结构, 所谓泛代数就是一个向量空间上可以有多种运算 [2]. 那么, 每一种运算用一个结构矩阵刻画, 于是一组结构矩阵就代表了一个特定的泛代数. 进一步的讨论可见 [5].

5 从多元多项式到非线性映射

设 $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. 那么, $\mathbb{R}^n$ 上的一个 (齐次) 线性函数 $f(x)$ 可表示成 $f(x) = cx$, 这里, $c = (c_1, \cdots, c_n)$ 为系数. 类此, $\mathbb{R}^n$ 上的 k 次齐次多项式 $P_k(x)$ 可表示成

$$P_k(x) = F_k x^k,$$

这里 $F_k \in \mathbb{R}^{n^k}$ 也称系数向量, 只是这时, F_k 不唯一. 例如, 设 $n = 3$, $k = 2$. 那么

$$x^2 = (x_1^2 \ x_1x_2 \ x_1x_3 \ x_2x_1 \ x_2^2 \ x_2x_3 \ x_3x_1 \ x_3x_2 \ x_3^2)^T.$$

设 $P_2(x) = 2x_1^2 + 3x_1x_3 - x_3^2$, 那么 $F_2 = (2\ 0\ 3\ 0\ 0\ 0\ 0\ 0\ -1)$, 或 $F_2 = (2\ 0\ 1\ 0\ 0\ 0\ 2\ 0\ -1)$, 等.

这个表示方法很容易推广到 $f : \mathbb{R}^n \to \mathbb{R}^k$. 例如, 一个 $\mathbb{R}^n$ 上的 k 次齐次的向量场 $\xi(x)$, 它可以表示成

$$\xi(x) = F_\xi x^k, \quad \text{这里} \ F_\xi \in \mathcal{M}_{n \times n^k}. \tag{27}$$

定义 2 *1. 设 $f(x) : \mathbb{R}^n \to \mathbb{R}$ 为一可微函数. 它的微分定义为*

$$Df(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n} \right). \tag{28}$$

2. 设 $M(x) \in \mathcal{M}_{p \times q}$ 为一可微函数矩阵, $x \in \mathbb{R}^n$. 那么, $M(x)$ 的微分, 记着 $DM(x) \in \mathcal{M}_{p \times nq}$, 可由将 $M(x)$ 的每个元素 $m_{i,j}$ 用它的微分 $\left(\frac{\partial dm_{i,j}}{\partial x_1}, \cdots, \frac{\partial m_{i,j}}{\partial x_n} \right)$ 来代替而生成. 即

$$DM(x) = \begin{bmatrix} \frac{\partial m_{11}(x)}{\partial x_1} & \cdots & \frac{\partial m_{11}(x)}{\partial x_n} & \cdots & \frac{\partial m_{1q}(x)}{\partial x_1} & \cdots & \frac{\partial m_{1q}(x)}{\partial x_n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial m_{p1}(x)}{\partial x_1} & \cdots & \frac{\partial m_{p1}(x)}{\partial x_n} & \cdots & \frac{\partial m_{pq}(x)}{\partial x_1} & \cdots & \frac{\partial m_{pq}(x)}{\partial x_n} \end{bmatrix}. \tag{29}$$

下面给一个一般的微分公式:

定理 3 设 $x=(x_1,\cdots,x_n)^T\in\mathbb{R}^n$. x^m 的微分满足以下公式:

$$D(x^{k+1})=\Phi_k^n x^k,\quad k\geq 0, \tag{30}$$

这里

$$\Phi_k^n=\sum_{s=0}^{k} I_{n^s}\otimes W_{[n^{k-s},n]}. \tag{31}$$

这些概念和公式很容易应用到一般非线性映射中去. 设 $f:\mathbb{R}^n\to\mathbb{R}^m$ 为一光滑映射. 我们可以将其 Taylor 展开, 得

$$f(x)=c_0+c_1x+c_2x^2+\cdots, \tag{32}$$

这里,

$$c_i\in\mathcal{M}_{m\times n^i},\quad i=0,1,\cdots.$$

如果要求 $f(x)$ 的微分, 则利用 (30) 可知

$$Df(x)=\sum_{i=0}^{\infty} c_{i+1}\Phi_i^n x^i. \tag{33}$$

这种方法可以很方便地用来处理非线性问题. 例如, 它在电力系统控制中的应用就是一个很好的例子 [8].

6 小结

本文分析了矩阵半张量积的基本原理: 它可以自动对参数 (或系数) 与变量找到对应关系. 这种自动寻配的能力, 加上它运算上的便捷, 使它成为一种处理多线性及非线性的有力工具. 由于经典矩阵方法只能处理线性或双线性问题, 半张量积大大扩大了矩阵的适用范围. 之后, 本文给出三类典型问题及其用半张量积处理的典型方法. 包括: (i) 有限集映射和有限集上的动态过程; (ii) 有限维代数 (泛代数) 的结构矩阵方法; (iii) 从多线性映射到非线性映射的半张量积表示及微分运算, 等. 它们为半张量积的运用提供了一个导向性的介绍. 作者相信, 半张量积会成为计算机时代的一个全新的既可用于理论分析, 又可用于计算的有力的数学工具.

参考文献

[1] D. Bates, D. Watts, Relative curvature measures of nonlinearity, *Journal of the Royal Statistical Society*, Series B (Methodological) Vol. 42, 1-25, 1980.

[2] S. Burris, H.P. Sankappanavar, *A Course in Universal Algebra*, Springer-Verlag, New York, 1981.

[3] 程代展, 齐洪胜, 《矩阵的半张量积－理论与应用》, 第 2 版, 科学出版社, 北京, 2011.

[4] D. Cheng, H. Qi, Z. Li, *Analysis and Control of Boolean Networks, A Semi-tensor Product Approach*, Springer, London, 2011.

[5] D. Cheng, H. Qi, Y. Zhao, *An Introduction to Semi-tensor Product and Its Applications*, World Scientific, Singapore, 2012. (to appear)

[6] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.

[7] 李文林,《数学史教程》, 高等教育出版社, 北京, 2000.

[8] 梅生伟, 刘锋, 薛安成,《电力系统暂态分析中的半张量积方法》, 清华大学出版社, 北京, 2010.

[9] C. Tsai, *Contributions to the Design and Analysis of Nonlinear Models*, Ph.D. thesis, Univ. of Minisota, 1983.

[10] 郑义, 赵建立, 李成允, 矩阵左半张量积的推广一泛张量积及其性质,《聊城大学学报》, Vol. 2, No. 1, 2010.

Principle and Range of Possible Applications of Sime-tensor Product of Matrices

Abstract: As a new matrix product, semi-tensor product of matrices has attracted more and more attention and participation from domestic and international academic society, and it has been applied to more and more research topics. The purpose of this paper is to analyze the fundamental principle of semi-tensor product, and to explain the reason for the emergence and existence of the semi-tensor product. Unlike the existing surveys, this paper does not intend to introduce what applications the semi-tensor product has been used to. Instead, we want to explore what kind of problems the semi-tensor product might be used according to the essence of semi-tensor product. Such observation could help us to dig out unknown possible further applications of semi-tensor product.

Keywords: Semi-tensor product of matrices, Boolean network, university algebra, multi-variable polynomials, matrix method of nonlinear problems.