Semi-tensor product of matrices and its application to Morgen's problem

CHENG Daizhan (程代展)

Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China (email: dcheng@iss03.iss.ac.cn)

Received October 10, 2000

Abstract This paper proposes a new matrix product, namely, semi-tensor product. It is a generalization of the conventional matrix product. Meanwhile, it is also closely related to Kronecker (tensor) product of matrices. The purpose of introducing this product is twofold: (i) treat multi-dimensional data; (ii) treat nonlinear problems in a linear way. Then the computer and numerical methods can be easily used for solving nonlinear problems. Properties and formulas are deduced. As an application, the Morgen's problem for control systems is formulated as a numerically solvable problem.

Keywords: swap matrix, left-semi-tensor product, mapping of matrices, decoupling matrix, Morgen's problem.

To treat a set of data, we know that if the data are of one dimension, they can be arranged as a vector, and data of two dimensions can be arranged into a matrix form. But what if the data have more than two dimensions? Computer scientists provide a solution: the data need not be arranged in an array of more than one dimension. In memory, they are arranged in a queue. But in programming, say in C-language, people use the so-called "pointer", "pointer-to-pointer", "pointer-to-pointer-to-pointer" etc. to distinguish different hierarchies of data. Motivated by this, the present paper proposes a new matrix multiplication called semi-tensor-product of matrices. Intuitively, what the semi-tensor-product is going to do is to provide a way to search the different levels of "pointers" automatically. Then it performs operations with respect to different layers of data mechanically.

From computational point of view one sees that a linear function can be realized as the inner-product of vectors; a quadratic form of matrices is a convenient way to investigate quadratic polynomials. Then can a higher degree polynomial be expressed in matrix form and calculated by matrix operations? If we can find such an operation, the tools in linear algebra may be used to treat nonlinear problems easily. In fact, semi-tensor-product, collaborated with tensor product, can handle arrays of data with dimensions higher that two. It is, therefore, a tool to treat higher degree polynomials. Then certain nonlinear problems can be treated in a linear way and manipulated in computer easily.

First, let us see three commonly used matrix multiplications: conventional matrix product, tensor product (Kronecker product)[1], and Hadamard product[2]. All of them satisfy two fundamental product rules: (i) associativity; (ii) distributivity.

Thus some sets of matrices are rings. Taking conventional scale product, they become algebras. A fundamental connection between conventional and tensor products is the following equation, which will be used from time to time:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \tag{0.1}$$

We refer to refs. [1, 2] for the other properties of these products.

1 Vector expression of matrices

Let $M \in M_{m \times n}$ be an $m \times n$ matrix. We denote its rows by $M^1, \cdots, M^m$, and its columns by $M_1, \cdots, M_n$, and then define two mappings V_r and V_c, which map $M_{m\times n} \to \mathbb{R}^{mn}$ as $V_r(M) = (M^1 M^2 \cdots M^m)^{\mathrm{T}}$ and $V_c(M) = (M_1^{\mathrm{T}} M_2^{\mathrm{T}} \cdots M_n^{\mathrm{T}})^{\mathrm{T}}$. The following equalities are the immediate consequences of the definition:

$$V_c(M) = V_r(M^{\mathrm{T}}); \quad V_r(M) = V_c(M^{\mathrm{T}}). \tag{1.1}$$

We define multi-indices as

Definition 1.1. Let N be a positive integer as $N = n_1 \times n_2 \times \cdots \times n_k$, with integers $n_j > 1$, $j = 1, \cdots, k$. Then we define a multi-index as

$$\mathrm{Id}(\lambda_1, \cdots, \lambda_k;\ n_1, \cdots, n_k) = \{(\lambda_1, \cdots, \lambda_k) \mid 1 \leqslant \lambda_j \leqslant n_j,\ \forall j\}.$$

Moreover, the elements in $\mathrm{Id}(\lambda_1, \cdots, \lambda_k; n_1, \cdots, n_k)$ are arranged in alphabetic order as

$$(\lambda_1, \cdots, \lambda_k) < (t_1, \cdots, t_k),$$

if there exists $0 \leqslant s < k$ such that $\lambda_1 = t_1, \cdots, \lambda_s = t_s$ and $\lambda_{s+1} < t_{s+1}$. Let $X = \{a_1, a_2, \cdots, a_N\}$ be a set of ordered N elements, $N = n_1 \times n_2 \times \cdots \times n_k$. Then we can use multi-indices $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_k\}$ as in Definition 1.1 to index the elements in X as $\{a_{(\lambda_{i_1}, \cdots, \lambda_{i_k})}\}$.

Example 1.1. Let $X = \{a_1, a_2, \cdots, a_{24}\}$. We index them by $(\lambda_1, \lambda_2, \lambda_3)$.

(i) Using $\mathrm{Id}(\lambda_1, \lambda_2, \lambda_3; 2, 3, 4)$, the elements are indexed as

$$a_{111}\, a_{112}\, a_{113}\, a_{114}\, a_{121}\, a_{122}\, a_{123}\, a_{124} \cdots a_{231}\, a_{232}\, a_{233}\, a_{234}.$$

(ii) Using $\mathrm{Id}(\lambda_2, \lambda_3, \lambda_1; 3, 4, 2)$, the elements are indexed as

$$a_{111}\, a_{211}\, a_{112}\, a_{212}\, a_{113}\, a_{213}\, a_{114}\, a_{214} \cdots a_{133}\, a_{233}\, a_{134}\, a_{234}.$$

Example 1.2. Let V be an n dimensional vector space, V^* its dual space. Let

$$\omega: \underbrace{V \times \cdots \times V}_{r} \times \underbrace{V^* \times \cdots \times V^*}_{s} \to \mathbb{R}$$

be a multi-linear mapping. (In fact, $\omega \in \boldsymbol{T}_s^r(V)$ is a tensor with covariant order r and contravariant order s. cf. refs. [3, 4].)

Let $\{d_1, \cdots, d_n\}$ be a basis of V with its dual basis $\{e_1, \cdots, e_n\}$. Then ω is determined by a set of data

$$\omega_{j_1 \cdots j_s}^{i_1 \cdots i_r} = \omega(d_{i_1}, \cdots, d_{i_r}, e_{j_1}, \cdots, e_{j_s}),\ i_1, \cdots, j_s = 1, \cdots, n.$$

Now we want to arrange this set of n^{r+s} elements into a matrix in such a way that the rows are arranged in the order of $\mathrm{Id}(i_1 \cdots i_r; \underbrace{n \cdots n}_{r})$ and the columns are arranged in the order of $\mathrm{Id}(j_1 \cdots j_s; \underbrace{n \cdots n}_{s})$. Then the matrix, M_ω, is

$$M_\omega = \begin{pmatrix} \omega_{1\cdots 11}^{1\cdots 11} & \omega_{1\cdots 11}^{1\cdots 12} & \cdots & \omega_{1\cdots 11}^{n\cdots nn} \\ \omega_{1\cdots 12}^{1\cdots 11} & \omega_{1\cdots 12}^{1\cdots 12} & \cdots & \omega_{1\cdots 12}^{n\cdots nn} \\ & & \cdots & \\ \omega_{n\cdots nn}^{1\cdots 11} & \omega_{n\cdots nn}^{1\cdots 12} & \cdots & \omega_{n\cdots nn}^{n\cdots nn} \end{pmatrix}. \tag{1.2}$$

Matrix M_ω is called the tensor array of the tensor ω. It will be used later.

Definition 1.2. A swap matrix, denoted by $W_{[m,n]}$, is an $mn \times mn$ matrix, defined as the following: Let its rows and columns be indexed by double index (i,j). The columns are indexed in the order of $\mathrm{Id}(i,j;m,n)$, while its rows be indexed in the order of $\mathrm{Id}(j,i;n,m)$. Then its element in the position $((I,J),(i,j))$ is set to

$$w_{(IJ),(ij)} = \delta_{i,j}^{I,J} = \begin{cases} 1, & I = i \text{ and } J = j, \\ 0, & \text{otherwise}. \end{cases}$$

The swap matrix plays a very important role in manipulating multi-indexed data.

Example 1.3. $W_{[2,3]}$ and $W_{[3,2]}$ are expressed as

$$W_{[2,3]} = \begin{array}{c} \begin{array}{cccccc} (11) & (12) & (13) & (21) & (22) & (23) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ (11) \\ (21) \\ (12) \\ (22) \\ (13) \\ (23) \end{array},$$

$$W_{[3,2]} = \begin{array}{c} \begin{array}{cccccc} (11) & (12) & (21) & (22) & (31) & (32) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ (11) \\ (21) \\ (31) \\ (12) \\ (22) \\ (32) \end{array}.$$

The following proposition follows from a straightforward computation:

Proposition 1.1. (i) The inverse and the transpose of a swap matrix are

$$W_{[n,m]} = W_{[m,n]}^{\mathrm{T}} = W_{[m,n]}^{-1}. \tag{1.3}$$

(ii) When $n = m$, (1.3) becomes

$$W_{[n,n]} = W_{[n,n]}^{\mathrm{T}} = W_{[n,n]}^{-1}. \tag{1.4}$$

$m = n$ is a particular and important case. To simplify the notation, we briefly denote $W_{[n,n]} = W_{[n]}$. It will be used through this paper.

Swap matrix can realize the swap between row and column stacking forms of a matrix:

Proposition 1.2. Let $A \in M_{m\times n}$. Then

$$\begin{cases} W_{[m,n]} V_r(A) = V_c(A), \\ W_{[n,m]} V_c(A) = V_r(A). \end{cases} \tag{1.5}$$

2 Left semi-tensor product of matrices

Definition 2.1. Let T be a row vector of dimension np, and X be a column vector with dimension p. Then we split T into p equal-size blocks as $T_1, \cdots, T_p$, which are $1 \times n$ rows. Define the left semi-tensor product, denoted by $\odot$, as

$$T \odot X = \sum_{i=1}^{p} T_i x_i \in \mathbb{R}^n. \tag{2.1}$$

Example 2.1. Let U, V and W be three vector spaces with dimension m, n and t respectively. $F \in L(U \times V \times W, \mathbb{R})$. Let $u_1, \cdots, u_m$; $v_1, \cdots, v_n$; $w_1, \cdots, w_t$ be the bases of U, V and W respectively. We define the structure constants as

$$s_{ijk} = F(u_i, v_j, w_k), \quad i = 1, \cdots, m; j = 1, \cdots, n; k = 1, \cdots, t.$$

Then we construct the tensor array, as an mnt dimensional row indexed by (i, j, k) in the order of $\mathrm{Id}(i, j, k; m, n, t)$ as

$$T = (s_{111} \cdots s_{11t} \cdots s_{1n1} \cdots s_{1nt} \cdots s_{mn1} \cdots s_{mnt}).$$

Then for $X \in U$, $Y \in V$ and $Z \in W$ we are ready to verify that

$$F(X, Y, Z) = T \odot X \odot Y \odot Z.$$

Now we are ready to give a general definition of the left semi-tensor product of matrices. We use $a \% b$ for the remainder of $\frac{a}{b}$.

Definition 2.2. (i) Let $X = (x_1, \cdots, x_s)$ be a row of vector and $Y = (y_1, \cdots, y_t)^{\mathrm{T}}$ a column of vector.

Case 1. If $s \% t = 0$, say, $s = t \times n$, then the left semi-tensor inner-product of X, Y is defined as a row of vector of dimension n as

$$\langle X, Y \rangle_L := \sum_{k=1}^{t} X^k y_k, \tag{2.2}$$

where $X = (X^1, \cdots, X^t)$, and each $X^i \in \mathbb{R}^n$, $i = 1, \cdots, t$.

Case 2. If $t \% s = 0$, say, $t = s \times n$, then the left semi-tensor inner-product of X, Y is defined as a column of vector of dimension n as

$$\langle X, Y \rangle_L := (\langle Y^{\mathrm{T}}, X^{\mathrm{T}} \rangle_L)^{\mathrm{T}}. \tag{2.3}$$

(ii) Let $M \in M_{m \times n}$ and $N \in M_{p \times q}$. If either n is a factor of p, or p is a factor of n, we define the left semi-tensor product of M and N, denoted by $C = M \odot N$, as the following: C consists of $m \times q$ blocks and each block is

$$C^{ij} = \langle M^i, N_j \rangle_L, \quad i = 1, \cdots, m, j = 1, \cdots, q.$$

Definition 2.3. Given a matrix A_{pq}, assume $p \% q = 0$ or $q \% p = 0$. Then we define A^n, $n > 0$ inductively as

$$\begin{cases} A^1 = A, \\ A^{k+1} = A^k \odot A, \quad k = 1, 2, \cdots. \end{cases}$$

Remark. It is easy to verify that the power is well defined. Moreover, if $p = sq$ (where s is a positive integer) the dimension of A^k is $(s^k q, q)$; if $q = sp$, the dimension of A^k is $(p, s^k p)$.

Example 2.2. (i) If X is a row vector or a column vector, then according to Definition 2.3, X^n is always defined. In fact, it is easy to see that when X, Y are two columns,

$$X \odot Y = X \otimes Y,$$

when X, Y are two rows,

$$X \odot Y = Y \otimes X.$$

Hence in either case

$$X^k = \underbrace{X \otimes \cdots \otimes X}_{k}.$$

(ii) Let X, Y be two columns and A, B be two matrices of proper dimensions. Then

$$(AX) \odot (BY) = (A \otimes B)(X \odot Y).$$

Hence

$$(AX)^k = \underbrace{(A \otimes \cdots \otimes A)}_{k} X^k.$$

(iii) Let X, Y be two rows and A, B be two matrices of proper dimensions. Then

$$(XA) \odot (YB) = (X \odot Y)(B \otimes A).$$

Hence

$$(XA)^k = X^k \underbrace{(A \otimes \cdots \otimes A)}_{k}.$$

(iv) Consider the set of k-th degree homogeneous polynomials of variables $X \in \mathbb{R}^n$, denoted by B_n^k. Under conventional polynomial addition and scalar multiplication, B_n^k is a vector space. It is obvious that X^k is a redundant basis. Hence any $P(x) \in B_n^k$ can be expressed as $P(x) = CX^k$, where the coefficients $C \in \mathbb{R}^{nk}$ are not unique.

Definition 2.4. Given $M_{m\times n}$ and $N_{p\times q}$.

(i) If $n\%p = 0$ we denote this situation by $M \succ N$, and if $p\%n = 0$ we denote it by $M \prec N$. $M \odot N$ is defined as either $M \succ N$ or $M \prec N$.

(ii) If $M \succ N$ and $n = tp$, to emphasize t we denote $M \succ_t N$; conversely, if $M \prec N$ and $nt = p$, we denote $M \prec_t N$.

(iii) For a finite sequence of matrices: $M_1, \cdots, M_n$ if either $M_1 \succ M_2 \succ \cdots \succ M_n$ or $M_1 \prec M_2 \prec \cdots \prec M_n$, the sequence will be called a proper chain.

Proposition 2.1. (i) Given $M_{m\times n}$ and $N_{p\times q}$ with proper dimensions n and p, then

$$M \odot N = \begin{pmatrix} M^1 \odot N_1 & \cdots & M^1 \odot N_q \\ & \vdots & \\ M^m \odot N_1 & \cdots & M^m \odot N_q \end{pmatrix} = \begin{pmatrix} M^1 \odot N \\ \vdots \\ M^m \odot N \end{pmatrix} = (M \odot N_1 \cdots M \odot N_q). \tag{2.4}$$

(ii) Assume $M \succ_t N$ (or $M \prec_t N$). Decompose M and N into blocks as

$$M = \begin{pmatrix} M^{11} & \cdots & M^{1s} \\ & \cdots & \\ M^{r1} & \cdots & M^{rs} \end{pmatrix}, N = \begin{pmatrix} N^{11} & \cdots & N^{1t} \\ & \cdots & \\ N^{s1} & \cdots & N^{st} \end{pmatrix}.$$

If $M^{ik} \succ_t N^{kj}$, $\forall\ i, j, k$ (or $M^{ik} \prec_t N^{kj}$, $\forall\ i, j, k$ respectively), then

$$M \odot N = \begin{pmatrix} C^{11} & \cdots & C^{1t} \\ & \cdots & \\ C^{r1} & \cdots & C^{rt} \end{pmatrix},$$

where

$$C^{ij} = \sum_{k=1}^{s} M^{ik} \odot N^{kj}.$$

(iii) If $M >_t N$ and $N >_s T$, then $M \odot N >_{ts} T$.

(iv) Let $M_1, \cdots, M_n$ be a proper chain. Then

$$M_1 \odot M_2 \odot \cdots \odot M_n$$

is well defined.

Now we are ready to prove the following fundamental properties:

Theorem 2.1. As long as the $\odot$ is defined, i.e. the matrices have suitable dimensions, the following are correct:

(i) Distributivity:

$$\begin{cases} F \odot (aG \pm bH) = aF \odot G \pm bF \odot H, \\ (aF \pm bG) \odot H = a F \odot H \pm bG \odot H, \quad a, b \in \mathbb{R}. \end{cases}$$

(ii) Associativity:

$$(F \odot G) \odot H = F \odot (G \odot H).$$

Proof. Part (i) is trivial. We prove part (ii).

First of all, we have to show that if F G and H have feasible dimensions for $(F \odot G) \odot H$ the dimensions are also feasible for $F \odot (G \odot H)$.

Case 1. $F \succ G$ and $G \succ H$: So the dimensions of F, G and H can be assumed to be $m \times np$, $p \times qr$ and $r \times s$ respectively.

Now the dimension of $F \odot G$ is $m \times nqr$. It is good for $(F \odot G) \odot H$. On the other hand the dimension of $G \odot H$ is $p \times qs$. It is good for $F \odot (G \odot H)$.

Case 2. $F \prec G$ and $G \prec H$: So the dimensions of F, G and H can be assumed as $m \times n$, $np \times q$ and $rq \times s$ respectively.

Now the dimension of $F \odot G$ is $mp \times q$. It is good for $(F \odot G) \odot H$. On the other hand the dimension of $G \odot H$ is $npr \times s$. It is good for $F \odot (G \odot H)$.

Case 3. $F \prec G$ and $G \succ H$. So the dimensions of F, G and H can be assumed as $m \times n$, $np \times qr$ and $r \times s$ respectively.

Now the dimension of $F \odot G$ is $mp \times qr$. It is good for $(F \odot G) \odot H$. On the other hand the dimension of $G \odot H$ is $np \times qs$. It is good for $F \odot (G \odot H)$.

Case 4. $F \succ G$ and $G \prec H$. So the dimensions of F, G and H can be assumed to be $m \times np$, $p \times q$ and $rq \times s$ respectively.

Now the dimension of $F \odot G$ is $m \times nq$. To make it feasible for $(F \odot G) \odot H$, we need:

Case 4.1. $(F \odot G) \succ H$, i.e. $n = n'r$. It is good for $F \odot (G \odot H)$.

Case 4.2. $(F \odot G) \prec H$, i.e. $r = nr'$. It is good for $F \odot (G \odot H)$.

The dimension of $G \odot H$ is $pr \times s$. To make it feasible for $(F \odot G) \odot H$, we need:

Case 4.3. $F \succ (G \odot H)$, i.e. $n = n'r$. It is good for $(F \odot G) \odot H$.

Case 4.4. $F \prec (G \odot H)$, i.e. $r = nr'$. It is good for $(F \odot G) \odot H$.

Next, we prove the associativity. We have to prove it case by case. But Cases 1—3 are similar, we prove only Case 1, i.e. $F \succ G$ and $G \succ H$.

Let $F_{m\times np}$, $G_{p\times qr}$ and $H_{r\times s}$ be given. Based on (2.4) we can, without loss of generality, assume $m = 1$ and $s = 1$. Then

$$F\odot G = (F_1\cdots F_p)\odot\begin{pmatrix} g_{11}^1 & \cdots & g_{1q}^1 & \cdots & g_{r1}^1 & \cdots & g_{rq}^1 \\ & \cdots & & \cdots & & \cdots & \\ g_{11}^p & \cdots & g_{1q}^p & \cdots & g_{r1}^p & \cdots & g_{rq}^p \end{pmatrix}$$

$$= \left(\sum_{i=1}^{p} F_i g_{11}^i \cdots \sum_{i=1}^{p} F_i g_{1q}^i \cdots \sum_{i=1}^{p} F_i g_{r1}^i \cdots \sum_{i=1}^{p} F_i g_{rq}^i \right).$$

Then we have

$$(F\odot G)\odot H = (F\odot G)\odot \begin{pmatrix} h_1 \\ \vdots \\ h_r \end{pmatrix} = \left(\sum_{j=1}^{r}\sum_{i=1}^{p} F_i g_{j1}^i h_j \cdots \sum_{j=1}^{r}\sum_{i=1}^{p} F_i g_{jq}^i h_j \right). \tag{2.5}$$

On the other hand

$$\begin{pmatrix} g_{11}^1 & \cdots & g_{1q}^1 & \cdots & g_{r1}^1 & \cdots & g_{rq}^1 \\ & \cdots & & \cdots & & \cdots & \\ g_{11}^p & \cdots & g_{1q}^p & \cdots & g_{r1}^p & \cdots & g_{rq}^p \end{pmatrix} \odot \begin{pmatrix} h_1 \\ \vdots \\ h_r \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{r} g_{j1}^1 h_j & \cdots & \sum_{j=1}^{r} g_{jq}^1 h_j \\ & \cdots & \\ \sum_{j=1}^{r} g_{j1}^p h_j & \cdots & \sum_{j=1}^{r} g_{jq}^p h_j \end{pmatrix}.$$

Then

$$F\odot (G\odot H) = (F_1\cdots F_p)\odot(G\odot H) = \left(\sum_{j=1}^{r}\sum_{i=1}^{p} F_i g_{j1}^i h_j \cdots \sum_{j=1}^{r}\sum_{i=1}^{p} F_i g_{jq}^i h_j \right),$$

which is the same as (2.5).

Cases 4.1—4.4 can be proved similarly.

3 Linear mapping on matrices

This section considers the matrix expression of a mapping of matrices. We start with examples:

Example 3.1. (i) (Lyapunov mapping) Given a square matrix, $A \in M_n$. Consider the following mapping L_A: $M_n \to M_n$, defined as

$$L_A(X) = AX + XA^{\mathrm{T}}. \tag{3.1}$$

It is well known that[5]: A is Hurwitz, if and only if for any negative definite matrix $Q < 0$, $L_A(X) = Q$ has positive definite solution. As a linear mapping on vector space M_n, L_A has a matrix expression as[6, 7]

$$M_L^c = A \otimes I + I \otimes A. \tag{3.2}$$

The precise meaning of the matrix expression is

$$V_c(L_A(X)) = M_L^c V_c(X). \tag{3.3}$$

(ii) (Symplectic mapping) A matric $X \in \mathrm{Sp}(2n, \mathbb{R})$, if and only if $JX + X^{\mathrm{T}}J = 0$, where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

In general, we may consider a generalized symplectic mapping

$$L_N(X) = NX + X^{\mathrm{T}}N.$$

It is proved in ref. [8] that for any N there exists a non-zero solution X for $L_N(X) = 0$.

In (3.3) since the matrix is expressed as a column stacking form, we use superscription, c, to indicate it. We can also have a matrix expression when the matrices are expressed in the row stacking form. That is

$$V_r(L_A(X)) = M_L^r V_r(X). \tag{3.4}$$

The following proposition shows that one matrix expression can be obtained from the other easily.

Proposition 3.1. Let $\rho \in L(M_{p\times q}, M_{m\times n})$. Then

$$\begin{cases} M_\rho^r = W_{[n,m]} M_\rho^c W_{[p,q]}, \\ M_\rho^c = W_{[m,n]} M_\rho^r W_{[q,p]}. \end{cases} \tag{3.5}$$

Particularly, if $\rho \in L(M_n)$, then (3.5) becomes

$$M_\rho^r = W_{[n]} M_\rho^c W_{[n]}, \qquad M_\rho^c = W_{[n]} M_\rho^r W_{[n]}.$$

According to (3.5), it is not necessary to consider both M^r and M^c. As a convention, when we say a matrix expression, it means M^c unless otherwise stated.

We consider a general linear mapping $\rho: M_{n\times p} \to M_{m\times q}$ as

$$Z \mapsto AZB + CZ^{\mathrm{T}}D, \tag{3.6}$$

where $A \in M_{m\times n}$, $B \in M_{p\times q}$, $C \in M_{m\times p}$ and $D \in M_{n\times q}$. For this general form we have:

Proposition 3.2. The matrix expression of (3.6) is

$$M^c = (B^{\mathrm{T}} \otimes A) + (D^{\mathrm{T}} \otimes C) W_{[p,n]}. \tag{3.7}$$

Proof. We first list matrix expressions for four basic linear mappings with precisely indicated dimensions.

$$Z \to AZ: \quad I_p \otimes A. \tag{3.8}$$

$$Z \to ZB: \quad B^{\mathrm{T}} \otimes I_n. \tag{3.9}$$

$$Z \to CZ^{\mathrm{T}}: (I_n \otimes C) W_{[p,n]}. \tag{3.10}$$

$$Z \to Z^{\mathrm{T}}D: \quad (D^{\mathrm{T}} \otimes I_p) W_{[p,n]}. \tag{3.11}$$

(3.8) and (3.9) are known[1, 2]. Now for (3.10) we have

$$\begin{aligned} V_c(CZ^{\mathrm{T}}) &= (I_n \otimes C) V_c(Z^{\mathrm{T}}) = (I_n \otimes C) V_r(Z) \\ &= (I_n \otimes C) W_{[p,n]} W_{[n,p]} V_r(Z) = (I_n \otimes C) W_{[p,n]} V_c(Z). \end{aligned}$$

(3.11) is proved as

$$\begin{aligned} V_c(Z^{\mathrm{T}}D) &= (D^{\mathrm{T}} \otimes I_p) V_c(Z^{\mathrm{T}}) = (D^{\mathrm{T}} \otimes I_p) V_r(Z) \\ &= (D^{\mathrm{T}} \otimes I_p) W_{[p,n]} W_{[n,p]} V_r(Z) = (D^{\mathrm{T}} \otimes I_p) W_{[p,n]} V_c(Z). \end{aligned}$$

Compounding the previous expressions provides the expressions for each term. Since both terms are linear mappings, a linear combination is allowed. Note that for the compound mappings the dimension should be adjusted. Then we have the following

$$M^c = (B^{\mathrm{T}} \otimes I_m)(I_p \otimes A) + (D^{\mathrm{T}} \otimes I_m)(I_n \otimes C) W_{[p,n]}.$$

(3.7) follows immediately.

(3.6) consists of only two terms with different types. By linear combination, (3.7) can be used for any finite term cases.

All the matrix expressions discussed above can be considered as certain particular cases of (3.7).

For convenience, we list the matrix expressions of some useful linear mappings of matrices in table 1.

Table 1 Matrix expression of certain linear mappings on matrices

Name	Symbol	$\rho(A)$	$M^c_{\rho(A)}$
Lyapunov	L_A	$Z \mapsto AZ + ZA^T$	$I \otimes A + A \otimes I$
General Lyapunov	L_{AB}	$Z \mapsto AZ + ZB$	$I \otimes A + B^T \otimes I$
Symplectic-type	S_A	$Z \mapsto AZ + Z^T A$	$I \otimes A + (A^T \otimes I) W$
Adjoint	ad_A	$Z \mapsto AZ - ZA$	$I \otimes A - A^T \otimes I$
Conjugate	Cj_A	$Z \mapsto AZA^{-1}$	$A^{-T} \otimes A$
Congruent	Cg_A	$Z \mapsto AZA^T$	$A \otimes A$

Based on previous expressions we may obtain some useful formulas. The following is a useful one:

Proposition 3.3. Let $A \in M_{m\times n}$ and $B \in M_{p\times q}$. Then

$$W_{[m,p]}(A \otimes B) W_{[q,n]} = (B \otimes A). \tag{3.12}$$

Particularly,

$$(I_p \otimes A) W_{[n,p]} = W_{[m,p]}(A \otimes I_p). \tag{3.13}$$

Proof. Let $Z \in M_{q\times n}$. Consider the expression $Z \mapsto AZ^T B^T$, and we realize it in the following two ways:

(i) $Z \mapsto ZA^T \mapsto BZA^T \mapsto (BZA^T)^T$: Then it is realized by $W_{[m,p]}(I_m \otimes B)(A \otimes I_q)$, which is $W_{[m,p]}(A \otimes B)$.

(ii) $Z \mapsto Z^T \mapsto AZ^T \mapsto AZ^T B^T$: Then it is realized by $(B \otimes I_m)(I_q \otimes A) W_{[n,q]}$, which is $(B \otimes A) W_{[n,q]}$.

It follows that

$$W_{[m,p]}(A \otimes B) = (B \otimes A) W_{[n,q]}.$$

Right-multiplying both sides by $W_{[q,n]}$, (3.12) follows.

4 Properties of left semi-tensor product

This section investigates the properties of the left semi-tensor product of matrices. At the same time, since many of the formulas are coupled with the Kronecker product, we are going to reveal the relationship between them.

Proposition 4.1. (i) Let $A \in M_{m\times np}$ and $B \in M_{p\times q}$. Then

$$A \odot B = A(B \otimes I_n). \tag{4.1}$$

(ii) Let $A \in M_{m\times n}$ and $B \in M_{np\times q}$. Then

$$A \odot B = (A \otimes I_p)B. \tag{4.2}$$

Proposition 4.2. Let the matrices A, X, Y be given with proper dimensions. Then

$$V_r(AX) = A \odot V_r(X). \tag{4.3}$$

$$V_c(YA) = A^{\mathrm{T}} \odot V_c(Y). \tag{4.4}$$

Proof. For (4.3), set $C = AX$ and denote the i-th row of $A$by A^i etc. According to (2.6), the i-th block of the right hand side is

$$A^i \odot V_r(X) = \langle (A^i, (x_{11} \cdots x_{1s} \cdots x_{n1} \cdots x_{ns})^{\mathrm{T}} \rangle_L$$

$$= \begin{pmatrix} \sum_{k=1}^{n} a_{ik} x_{k1} \\ \cdots \\ \sum_{k=1}^{n} a_{ik} x_{ks} \end{pmatrix} = (C^i)^{\mathrm{T}}.$$

Eq. (4.3) follows. Using (1.1) to (4.3), one sees (4.4) easily.

Note that the form of (4.3) is very much like that of linear mapping on a vector space (such as $\mathbb{R}^n$ etc.). (4.3) can be used to obtain the matrix expression of a polynomial of a matrix. The following is straightforward.

Corollary 4.1. Let X be a square matrix and $p(x)$ be a polynomial. Then $p(x)$ can be expressed as $p(x) = q(x)x + p_0$, and

$$V_r(p(X)) = q(X) \odot V_r(X) + p_0 \odot V_r(I). \tag{4.5}$$

Proposition 4.3. Let $A \in M_{m\times n}$. Then

$$W_{[m,q]} \odot A \odot W_{[q,n]} = I_q \otimes A. \tag{4.6}$$

Proof. Choose $X \in M_{n\times q}$. According to (4.3) we have

$$V_r(AX) = A \odot V_r(X) = A \odot W_{[q,n]} V_c(X).$$

Left-multiplying both sides by $W_{[m,q]}$ yields

$$V_c(AX) = (W_{[m,q]} \odot A \odot W_{[q,n]}) V_c(X). \tag{4.7}$$

Comparing (3.8) with (4.7) yields (4.6).

Corollary 4.2. (i) Let $A \in M_{p\times n}$, $X \in M_{m\times p}$. Then

$$V_r(XA) = (I_m \otimes A^{\mathrm{T}}) \odot V_r(X). \tag{4.8}$$

(ii) Let $A \in M_{n\times p}$, $Y \in M_{p\times m}$. Then

$$V_c(AY) = (I_m \otimes A) \odot V_c(Y). \tag{4.9}$$

Proof. (i) Using (1.5) and (4.4),

$$V_r(XA) = W_{[n,m]} \odot V_c(XA) = W_{[n,m]} \odot A^{\mathrm{T}} \odot V_c(X) = W_{[n,m]} \odot A^{\mathrm{T}} \odot W_{[m,p]} \odot V_r(X).$$

Using (4.6),

$$W_{[n,m]} \odot A^{\mathrm{T}} \odot W_{[m,p]} \odot V_r(X) = (I_m \otimes A^{\mathrm{T}}) \odot V_r(X).$$

(ii) Using (1.1) we have

$$V_c(AY) = V_r(Y^{\mathrm{T}} A^{\mathrm{T}}) = (I_m \otimes A) \odot V_r(Y^{\mathrm{T}}) = (I_m \otimes A) \odot V_c(Y).$$

Note that the conventional product is a special case of left semi-tensor product, so anytime we can convert conventional product to left semi-tensor product.

Roughly speaking, a swap matrix can also "swap" a matrix with a vector. The following formulae are convenient in some cases:

Proposition 4.4. (i) Let Z be a row of dimension t and $A \in M_{m\times n}$. Then

$$Z\odot W_{[m,t]}\odot A = A\odot Z\odot W_{[n,t]} = A\otimes Z. \tag{4.10}$$

(ii) Let Y be a column of dimension t and A as above. Then

$$A\odot W_{[t,n]}\odot Y = W_{[t,m]}\odot Y\odot A = A\otimes Y. \tag{4.11}$$

Proof. (i) It is easy to see that

$$W_{[m,t]} = \begin{pmatrix} I_m\otimes\delta^1 \\ \cdots \\ I_m\otimes\delta^t \end{pmatrix}, \tag{4.12}$$

where $\delta^i = (0\cdots \underset{i\text{-th}}{1}\cdots 0)\in\mathbb{R}^t$. Then we have

$$Z\odot W_{[m,t]} = \sum_{j=1}^{t} z_j I_m\otimes\delta_j = I_m\otimes Z. \tag{4.13}$$

Using (4.1), we have

$$Z\odot W_{[m,t]}\odot A = (I_m\otimes Z)\odot A = (I_m\otimes Z)(A\otimes I_t) = A\otimes Z.$$

Similarly, we have

$$A\odot Z\odot W_{[n,t]} = A\odot(I_n\otimes Z) = (A\otimes I_1)(I_n\otimes Z) = A\otimes Z.$$

(ii) In (4.10) replace A by A^{T}, Z by Y^{T} and then take transpose on both sides. Note that $W^{\mathrm{T}}_{[m,n]} = W_{[n,m]}$. (4.11) follows.

Remark. (4.12) is a convenient expression of the swap matrix in some calculations. It has a dual form as the following, which is also useful:

$$W_{[m,t]} = (I_t\otimes\delta_1\cdots I_t\otimes\delta_m), \tag{4.14}$$

where $\delta_i = (0\cdots \underset{i\text{-th}}{1}\cdots 0)^{\mathrm{T}}\in\mathbb{R}^m$.

The following matrix-vector swapping formulas are very convenient in practical use.

Corollary 4.3. Assume $A\in M_{m\times n}$. (i) Let $Z\in\mathbb{R}^t$ be a row vector. Then

$$A\odot Z = Z\odot W_{[m,t]}\odot A\odot W_{[t,n]} = Z\odot(I_t\otimes A). \tag{4.15}$$

(ii) Let $Z\in\mathbb{R}^t$ be a column vector. Then

$$Z\odot A = W_{[m,t]}\odot A\odot W_{[t,n]}\odot Z = (I_t\otimes A)\odot Z. \tag{4.16}$$

(iii) Let $X\in\mathbb{R}^t$ be a column, $Y\in\mathbb{R}^s$ be a row. Then

$$XY = Y\odot W_{[t,s]}\odot X. \tag{4.17}$$

Proposition 4.1 shows that the left semi-tensor product of two matrices can easily be expressed via Kronecker product. Using Corollary 4.3, we can also express Kronecker product of two matrices by the left semi-tensor product.

Proposition 4.5. Let $A\in M_{m\times n}$ and $B\in M_{s\times t}$. Then

$$A\otimes B = W_{[s,m]}\odot B\odot W_{[m,t]}\odot A = (I_m\otimes B)\odot A. \tag{4.18}$$

Proposition 4.6. (i) Let M be a $m \times pn$ matrix. Then

$$M \odot I_n = M. \tag{4.19}$$

(ii) Let M be an $m \times n$ matrix. Then

$$M \odot I_{pn} = M \otimes I_p. \tag{4.20}$$

(iii) Let M be a $pm \times n$ matrix. Then

$$I_p \odot M = M. \tag{4.21}$$

(iv) Let M be an $m \times n$ matrix. Then

$$I_{pm} \odot M = M \otimes I_p. \tag{4.22}$$

Proposition 4.7. Given two invertible matrices A and B with proper dimensions, such that $A \odot B$ is defined, then

(i) $A \odot B$ and $B \odot A$ have the same characteristic function.

(ii) $\mathrm{Trace}(A \odot B) = \mathrm{Trace}(B \odot A)$.

(iii) If either A or B is invertible, then $A \odot B \sim B \odot A$.

(iv) If both A and B are invertible, then

$$(A \odot B)^{-1} = B^{-1} \odot A^{-1}. \tag{4.23}$$

5 Tensor form of polynomials

Let $X = (x_1, \cdots, x_n)$ be a set of coordinates in $\mathbb{R}^n$. Denote by B_n^k the set of k-th order homogeneous polynomials of X. For convenience, set $B_n^0 = \mathbb{R}$.

It was pointed out before that X^k is a basis of $B_n{}^k$. So $f(x) \in B_n^k$ can be expressed as

$$f(x) = F \odot X^k, \text{ where } F \in M_{1 \times n^k}. \tag{5.1}$$

Since X^k is a redundant basis, the expression of F is not unique. Similar to the expression of a quadratic form, $f(x)$, in linear algebra, we define a symmetric expression.

Definition 5.1. Let $f(x) \in B_n^k$ and F be a representation of $f(x)$ as in (5.1). F is called a symmetric representation if under multi-index $(i_1, \cdots, i_k)$ in the order of $\mathrm{Id}(i_1, \cdots, i_k; n, \cdots, n)$

$$F_{i_{\sigma(1)}, \cdots, i_{\sigma(k)}} = F_{i_1, \cdots, i_k}, \quad \forall \sigma \in S_k,$$

where S_k stands for symmetric group of k elements, i.e. σ is a permutation of k elements.

Given an expression (5.1) of $f(x)$, the symmetric expression of it can be obtained as

$$\tilde{F}_{i_1, \cdots, i_k} = \frac{1}{k!} \sum_{\sigma \in S_n} F_{i_{\sigma(1)}, \cdots, i_{\sigma(k)}}. \tag{5.2}$$

Let $P(x)$, $x \in \mathbb{R}^n$ be the vector space of polynomials. Then

$$P = B_n^0 \oplus B_n^1 \oplus B_n^2 \oplus \cdots. \tag{5.3}$$

Proposition 5.1. Let $P(x) \in B_n^p$ and $Q(x) \in B_n^q$, and the corresponding matrix expressions be M_P and M_Q respectively. Then M_{PQ}, the matrix expression of the product $P(x)Q(x) \in B_n^{p+q}$, is

$$M_{PQ} = M_P \odot M_Q. \tag{5.4}$$

Proposition 5.1 shows that the advantage of the above expression is that under tensor product the products between factors are associative. It is more convenient and natural in application. For instance, we may factorize the multi-variable polynomials in a similar way as what we do for single variable polynomials, e.g.

$$F_3 \odot X^3 + F_5 \odot X^5 = (F_3 + F_5 \odot X^2) \odot X^3,$$

$$(F_1 \odot X + F_2 \odot X^2)(F_1 \odot X - F_2 \odot X^2) = F_1{}^2 \odot X^2 - F_2{}^2 \odot X^4,$$

etc.

The following proposition is an immediate consequence.

Proposition 5.2. Let $M_P \in \mathbb{R}^{n^p}$ and $M_Q \in \mathbb{R}^{n^q}$ be two rows. Then

$$M_P \odot M_Q \odot X^{p+q} = M_Q \odot M_P \odot X^{p+q}. \tag{5.5}$$

Note that since X^k is a redundant basis of B_n^k, (5.5) does not imply that $M_P \odot M_Q = M_Q \odot M_P$.

Next, we consider the multi-dimensional polynomial mappings. An m-dimensional polynomial of degree k with n arguments can be expressed as

$$P(x) = A_0 + A_1 X + A_2 X^2 + \cdots + A_k X^k, \qquad X \in \mathbb{R}^n,\ P(X) \in \mathbb{R}^m,$$

where $A_j \in M_{m \times n^j}$, $j = 0, \cdots, k$.

We consider the tensor product of two polynomial mappings.

Proposition 5.3. Let $P(x) = A_0 + A_1 X + A_2 X^2 + \cdots + A_p X^p$, $Q(x) = B_0 + B_1 X + B_2 X^2 + \cdots + B_q X^q$. Then

$$P(X) \odot Q(X) = \sum_{i=0}^{p+q} \sum_{j=0}^{i} (A_j \otimes B_{i-j}) X^i. \tag{5.6}$$

6 Formulation on Morgen's problem

As an application of the semi-tensor product of the matrices to practical problems, the rest of this paper will be devoted to the numerical formulation of the Morgen's problem.

Consider a linear system

$$\begin{cases} \dot{x} = Ax + Bu, & x \in \mathbb{R}^n,\ u \in \mathbb{R}^m, \\ y = Cx, & y \in \mathbb{R}^p. \end{cases} \tag{6.1}$$

Morgen's problem[9] asked when we can find a control $u = Kx + Hv$, with $v \in \mathbb{R}^p$ such that the closed-loop system is decoupled, i.e. each v_i controls y_i and does not affect y_j, $j \neq i$. Rigorously speaking, the transfer matrix is non-singular and diagonal. Denote by $\rho_1, \cdots, \rho_p$ the relative degrees of the outputs. Then when $m = p$ the answer to Morgen's problem is:

Theorem 6.1[10]. When $m = p$, the Morgen's problem is solvable if and only if the decoupling matrix

$$D = \begin{pmatrix} c_1 A^{\rho_1 - 1} B \\ \vdots \\ c_p A^{\rho_p - 1} B \end{pmatrix} \tag{6.2}$$

is non-singular. As $m > p$, the problem has been discussed over 35 years. It was claimed to be solved several times, e.g. refs. [11—13]. But then some counter-examples were discovered. So far, it is still an open problem. The problem was partly solved in refs. [14, 15] etc. In this paper, we will convert the problem to a numerically solvable problem.

From Theorem 6.1, the following is obvious:

Lemma 6.1. The Morgen's problem is solvable if and only if there exist $K \in M_{m\times n}$, $H \in M_{m\times p}$ and $1 \leqslant \rho_i \leqslant n$, $i = 1,\cdots,p$, such that

$$c_i(A + BK)^{t_i}BH = 0, \quad t_i = 0,\cdots,\rho_i - 2, i = 1,\cdots,p, \tag{6.3}$$

and

$$D = \begin{pmatrix} c_1(A + BK)^{\rho_1 - 1}BH \\ \vdots \\ c_p(A + BK)^{\rho_p - 1}BH \end{pmatrix} \tag{6.4}$$

is non-singular.

Let

$$W(K) = \begin{pmatrix} c_1B \\ c_1(A + BK)B \\ \vdots \\ c_1(A + BK)^{\rho_1 - 2}B \\ \vdots \\ c_p(A + BK)^{\rho_p - 2}B \end{pmatrix}; \quad T(K) = \begin{pmatrix} c_1(A + BK)^{\rho_1 - 1}B \\ \vdots \\ c_p(A + BK)^{\rho_p - 1}B \end{pmatrix}.$$

Then (6.3) becomes

$$W(K)H = 0, \tag{6.5}$$

and (6.4) becomes

$$D = T(K)H. \tag{6.6}$$

Since $1 \leqslant \rho_i \leqslant n$, $i = 1,\cdots,p$, we may consider the solvability of Morgen's problem for fixed ρ_i, $i = 1,\cdots,p$, because only finite cases (n^p) have to be verified. For statement ease, in the rest of this paper we consider the solvability of the Morgen's problem under a set of fixed ρ_i unless otherwise stated.

Lemma 6.2. The Morgen's problem is solvable if and only if there exists a K such that C1:

$$\mathrm{Im}(T^{\mathrm{T}}(K)) \cap \mathrm{Im}(W^{\mathrm{T}}(K)) = 0, \tag{6.7}$$

and C2: $T(K)$ has full row rank.

Proof. We show the following are equivalent:

(I) There exists H such that $T(K)H$ is non-singular and $W(K)H = 0$;

(II) $T(K)((W(K)^{\mathrm{T}})^{\perp}) = \mathbb{R}^p$;

(III) $[(T^{\mathrm{T}}(K))^{-1}(W(K)^{\mathrm{T}})]^{\perp} = \mathbb{R}^p$;

(IV) $(T^{\mathrm{T}}(K))^{-1}(W(K)^{\mathrm{T}}) = 0$;

(V) C1 and C2,

where $T(K)$ and $(T^{\mathrm{T}}(K))^{-1}$ are considered as a functional mapping[17].

(I)→(II): If $\dim(T(k)((W(K)^{\mathrm{T}})^{\perp})) < p$, since $\mathrm{Im}(H) \subset (W(K)^{\mathrm{T}})^{\perp}$, then $\mathrm{rank}(T(K)H) < p$ which is a contradiction. (II)→(I): Choose p vectors $h_i \in (W(K)^{\mathrm{T}})^{\perp}$, such that $T(K)\mathrm{Im}(h_1, \cdots, h_p) = \mathbb{R}^p$. Then set $H = (h_1, \cdots, h_p)$.

(II)↔(III): From page 23 of ref. [17]. (III)↔(IV) is obvious. (IV)↔(V): It is easy to verify that both (IV) and (V) are equivalent to the following statement: If $Y \in \mathbb{R}^p$ and $T^{\mathrm{T}}(K)Y \in \mathrm{Im}((W(K))^{\mathrm{T}})$, then $Y = 0$.

The following theorem is an immediate consequence of the above lemma.

Theorem 6.3. Morgen's problem is solvable (for fixed ρ_i) if and only if there exists a K_0 such that

$$\mathrm{rank}\begin{pmatrix} T(K_0) \\ W(K_0) \end{pmatrix} = p + \mathrm{rank}(W(K_0)). \tag{6.8}$$

Definition 6.1. Let $A(K)$ be a matrix with the entries $a_{ij}(K)$ as polynomials of K, where $K \in M_{m\times n}$. The essential rank of $A(K)$, denoted by $\mathrm{rank}_e(A(K))$ is defined as

$$\mathrm{rank}_e(A(K)) = \max_{K\in M_{m\times n}} \mathrm{rank}(A(K)).$$

Now (under fixed ρ_i) denote

$$\mathrm{rank}_e(T(K)) = t; \quad \mathrm{rank}_e(W(K)) = s, \quad \mathrm{rank}_e\begin{pmatrix} T(K) \\ W(K) \end{pmatrix} = q.$$

Since the essential rank is easily computable, the following corollary, containing a sufficient condition and a necessary condition, is convenient in some cases.

Corollary 6.1. (i) Morgen's problem is solvable if $q = p + s$. (ii) Morgen's problem is not solvable if $t < p$.

Note that since the computation of the essential rank is a mechanical procedure, Corollary 6.1 is easily verifiable.

7 Numerical method of Morgen's problem

To begin with, we have to calculate $T(K)$ and $W(K)$. Denote $Z = V_r(K) \in \mathbb{R}^{mn}$. We first express $T(K)$ and $W(K)$ as the polynomial matrices of Z. The product of two matrices can be expressed by a semi-tensor product of matrices:

Lemma 7.1. Given a matrix $A \in M_{n\times m}$.

(i) If $X \in \mathbb{R}^n$ is a row, then

$$XA = V_r^{\mathrm{T}}(A) \odot X^{\mathrm{T}}. \tag{7.1}$$

(ii) If $Y \in M_{p\times n}$, then

$$YA = (I_p \otimes V_r^{\mathrm{T}}(A)) \odot V_r(Y). \tag{7.2}$$

Proof. (7.1) can be obtained from a straightforward computation. Using (7.1) to each row of Y, (7.2) follows.

Now we expand $(A + BK)^t$ in the following way:

$$(A + BK)^t = \sum_{i=0}^{2^t-1} P_i(A, BK),$$

where P_i is a product of t elements, which are either A or BK. It is constructed in the following way: Convert i into a binary form, and replace each "0" by "A" and "1" by "BK". Collecting terms with the same number of "K", we can easily get the following expression:

$$c_k(A + BK)^t B = \sum_{i=0}^{t} \sum_{j=1}^{T_i} S_0^{ij} K S_1^{ij} K \cdots S_{t-1}^{ij} K S_t^{ij}, \tag{7.3}$$

where $T_i = \begin{pmatrix} t \\ i \end{pmatrix}$. Using Lemma 7.1 and eq. (4.16), (7.3) can be expressed as

$$\begin{aligned} c_k(A + BK)^t B &= \sum_{i=0}^{t} \sum_{j=1}^{T_i} S_0^{ij} \odot (I_m \otimes V_r^{\mathrm{T}}(S_1^{ij})) \odot Z \odot \cdots \\ &\odot (I_m \otimes V_r^{\mathrm{T}}(S_t^{ij})) \odot Z \\ &= \sum_{i=0}^{t} \sum_{j=1}^{T_i} S_0^{ij} \odot (I_m \otimes V_r^{\mathrm{T}}(S_1^{ij})) \odot (I_{m^2 n} \otimes V_r^{\mathrm{T}}(S_2^{ij})) \odot \\ &\cdots \odot (I_{m^t n^{t-1}} \otimes V_r^{\mathrm{T}}(S_t^{ij})) \odot Z^t, \quad k = 1, \cdots, p. \end{aligned} \tag{7.4}$$

Using (7.4), we can easily express $W(K)$ and $T(K)$ in the standard polynomial form as

$$\begin{aligned} W(K) &= W_0 + W_1 \odot Z + \cdots + W_{L-1} \odot Z^{l-1} \in M_{d \times m}, \quad d = \sum_{i=1}^{p} \rho_i - p, \\ T(K) &= T_0 + T_1 \odot Z + \cdots + T_L \odot Z^l \in M_{p \times m}. \end{aligned} \tag{7.5}$$

where $l = \max\{\rho_i - 1 \mid i = 1, \cdots, p\}$.

Let W^s be the set of s rows of $W(K)$. Then the size

$$|W^s| = \frac{d!}{s!(d-s)!}.$$

Now the Morgen's problem can be formulated as:

Proposition 7.1. The Morgen's problem is solvable if and only if there exists $1 \leqslant s \leqslant m - p + 1$ such that

$$R(Z) := \sum_{L \in W^s} \det(L(Z) L^{\mathrm{T}}(Z)) = 0, \tag{7.6}$$

and

$$J(Z) := \sum_{L \in W^{s-1}} \det\left(\begin{pmatrix} T(Z) \\ L(Z) \end{pmatrix} (T^{\mathrm{T}}(Z) L^{\mathrm{T}}(Z)) \right) > 0 \tag{7.7}$$

have a solution Z.

Proof. If (7.6) and (7.7) have a solution $Z = V_r(K)$, then (7.6) implies that $\operatorname{rank}(W(K)) < s$, and (7.7) implies

$$\mathrm{rank}\begin{pmatrix} T(K) \\ W(K) \end{pmatrix} = p + s - 1.$$

Then the conclusion is an immediate consequence of Theorem 6.3.

Now the Morgen's problem becomes a numerical problem: For fixed $1 \leqslant \rho_i \leqslant n$, $i = 1, \cdots, p$ and $1 \leqslant s \leqslant m - p + 1$, to solve (7.6), (7.7). Since there are only finite cases, if only the problem is solvable in each case, the Morgen's problem is solvable. Many numerical methods can be implemented to solve this problem.

For instance, it can be converted to standard Wu's problem[16]: Does polynomial $R(Z) = 0$ imply polynomial $J(Z) = 0$? "Yes" means the Morgen's problem is not solvable; and "No" means it is solvable. Wu's method can be used to solve this problem.

We can also formulate the problem as an optimization problem:

$$\max_{R(Z)=0} J(Z).$$

If the maximum value is zero, the Morgen's problem is not solvable. Otherwise, it is solvable.

Note that if each element of a square matrix, $A(Z)$, is expressed as a standard polynomial form as $a_0 + a_1 \odot Z + \cdots + a_l \odot Z^l$, to calculate $\det(A(Z))$ as a polynomial is a straightforward computation. Therefore, to get (7.6) and (7.7), we have only to calculate the following product. Let

$A = A_0 + A_1 \odot Z + \cdots + A_s \odot Z^s \in M_{m\times n}$; $B = B_0 + B_1 \odot Z + \cdots + B_t \odot Z^t \in M_{p\times n}$.

Then

$$AB^{\mathrm{T}} = \sum_{i=0}^{s}\sum_{j=0}^{t} A_i \odot Z^i \odot (Z^{\mathrm{T}})^j \odot B_j^{\mathrm{T}}.$$

Using (7.1)

$$(Z^{\mathrm{T}})^j = (Z^{\mathrm{T}})^j I_{n^j} = V_r^{\mathrm{T}}(I_{n^j}) \odot Z^j.$$

Using (4.16)

$$Z^i \odot V_r^{\mathrm{T}}(I_{n^j}) = (I_{n^i} \otimes V_r^{\mathrm{T}}(I_{n^j})) \odot Z^i,$$
$$Z^{i+j} \odot B_j^{\mathrm{T}} = (I_{n^{i+j}} \otimes B_j^{\mathrm{T}}) \odot Z^{i+j}.$$

Using them, we have

$$AB^{\mathrm{T}} = \sum_{i=0}^{s}\sum_{j=0}^{t} A_i \odot (I_{n^i} \otimes V_r^{\mathrm{T}}(I_{n^j})) \odot (I_{n^{i+j}} \otimes B_j^{\mathrm{T}}) \odot Z^{i+j}. \tag{7.8}$$

Summarizing the above argument we propose the following procedure for solving Morgen's problem:

1. For $\rho_1, \cdots, \rho_p = 1, \cdots, n$, using (7.3), (7.4) to express $T(K)$ and $W(K)$ into stardard polynomial matrix form (7.5).

2. For $s = 1, \cdots, m - p + 1$ and each $L \in W^s$, using (7.8) to calculate

$$L(Z)L^{\mathrm{T}}(Z), \text{ and } \begin{pmatrix} T(Z) \\ L(Z) \end{pmatrix} (T^{\mathrm{T}}(Z)\ L^{\mathrm{T}}(Z)).$$

3. Calculate $R(Z)$ and $J(Z)$ by (7.6) and (7.7) respectively.

4. Solve the numerical problem

$$\begin{cases} R(Z) = 0, \\ J(Z) > 0, \end{cases} \tag{7.9}$$

where both $R(Z)$ and $J(Z)$ are only polynomials.

8 Conclusion

A new matrix product, called the left-semi-tensor product, was introduced. As an auxiliary tool, the swap matrix was also introduced. The comparison between the new product with both conventional product and the Kronecker (tensor) product was presented. Certain properties and formulas were deduced. As an application, the numerical solution of the Morgen's problem was considered. A numerically verifiable necessary and sufficient condition was provided.

In fact, the new tool has been applied to several other problems, such as control Routh array, Carleman linearization, Hamiltonian realization of dynamic systems, etc. In one word, the new product is a useful tool.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. G69774008, G59837270), National 973 Project (Grant No. G1998020308) and National Key Project of China.

References

1. Huang L., Linear Algebra in Systems and Control Theory (in Chinese), Beijing: Science Press, 1984.
2. Zhang, F., Matrix Theory, Basic Results and Techniques, New York: Springer-Verlag, 1999.
3. Sokolnikoff, I. S., Tensor Analysis, Theory and Applications to Geometry and Mechanics of Continua, 2nd ed., New York: John Wiley & Sons, Inc., 1964.
4. Boothby, W. M., An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd ed., New York: Academic Press, 1986.
5. Willems, J. L., Stability Theory of Dynamical Systems, New York: John Wiley & Sons, Inc., 1970.
6. Ooba, T., Funahashi, Y., Two conditions concerning common quadratic Lyapunov functions for linear systems, IEEE Trans. Automat. Contr., 1997, 42(5): 719—721.
7. Ooba, T., Funahashi, Y., Stability robustness for linear state space models——a Lyapunov mapping approach, Sys. Contr. Lett, 1997, 29. 191—196.
8. Cheng, D., Xue, W., Huang, J., On general Hamiltonian Systems, Proc. of ICARCV'98, Singapore, 1998, 185—189.
9. Morgan, B. S., The synthesis of linear multivariable systems by state feedback, JACC, 1964, 64: 468—472.
10. Falb, P. L., Wolovich, W. A., Decoupling and synthesis of multivariable control systems, IEEE Trans, Aut. Contr., 1967, 12: 651—668.
11. Wolovich, W. A., Linear Multivariable System, New York: Springer-Verlag, 1974, 296.
12. Suda, M., Unahashi, K., Decoupling of nonsquare systems——a necessary and sufficient condition in terms of infinite zeros, Proc. 9th IFAC World Conf., Kyoto, 1984, Vol. 8, 88—93.
13. Descusse, J., Lafay, J. F., Malabre, M., Solution to Morgan's problem, IEEE Trans. Aut. Contr., 1988, 33: 732—739.
14. Chen, S., Cao, L., On constrained Morgen's problem, Science in China (in Chinese), Ser. A, 1996, 26(6): 513—523.
15. Xu, K., On a class of non-square Morgen's problem, Science in China (in Chinese), Ser. A, 1996, 26(4): 694—703.
16. Wu, W., On Mechanization of Mathematics (Partly in Chinese), Jinan: Shandong Educational Press, 1995.
17. Wonham, W. M., Linear Multivariable Control, A Geometric Approach, Berlin: Springer-Verlag, 1974.