Some Applications on Tensor Products

KyawZaw[1]

Abstract

Tensor product is a nice tool in Mathematics. In this paper, we discuss properties of tensor products. We demonstrate the applications in some problems related with classical mechanics, quantum mechanics, graph theory and image processing by using tensor products.

Keywords: Tensor; Matrix; Mechanics; Graph.

Introduction

In linear algebra there are many types of products:

The scalar product: $V \times F \to V$.

The dot product: $R^n \times R^n \to R$.

The cross product: $R^3 \times R^3 \to R^3$.

Matrix products: $M_{m\times k} \times M_{k\times n} \to M_{m\times n}$.

Note that the three vector spaces involved are not necessarily the same. The above examples have a common fact that the product is a bilinear map. The tensor product is just another example of a product like this.

1. Axiomatic Notions of Tensor Product

1.1 Definition

Let V_1, V_2 be vector spaces over a field F. A pair (Y, μ), where Y is a vector space over F and $\mu : V_1 \times V_2 \to Y$ is a bilinear map, is called the **tensor product** of V_1 and V_2 if the following condition holds:

Whenever β_1 is a basis for V_1 and β_2 is a basis for V_2, then

$$\mu(\beta_1 \times \beta_2) = \{\mu(x_1, x_2) \mid x_1 \in \beta_1, x_2 \in \beta_2\} \text{ is a basis for } Y .$$

We write $V_1 \otimes V_2$ for the vector space Y, and $x_1 \otimes x_2$ for $\mu(x_1, x_2)$.

1.2 Example.

(i) If V is any vector space over F, then $V \otimes F = V$. In this case, $\otimes$ is just scalar multiplication.

(ii) An inner product structure on a real vector space V is a bilinear map

$$F = \langle .,. \rangle : V \times V \to R$$

satisfying some additional conditions (positivity, definiteness, symmetry).

[1]Associate Professor, Department of Mathematics, University of Yangon

(iii)Let V = F^{n}_{row} , and W = F^{m}_{col} . Then V ⊗ W = $M_{m\times n}$(F), with the product defined to be v ⊗ w = wv, the matrix product of a column and a row vector.

Note: We can also say that W ⊗V = $M_{m\times n}$(F), with the product defined to be w⊗v = wv. From this, it looks like W ⊗ V and V ⊗ W are the same space.

1.3 Properties.

The vector space V_1⊗ V_2 and the map μ satisfy the following fundamental properties:

1. The map μ is bilinear. We sometimes call μ vector multiplication. Call any vector in the image of μ (any element in V_1⊗ V_2 of the form v ⊗ w) a **pure tensor.**

By bilinearityof μ, we see that

$$(av + bv) \otimes w = \mu(av + bv', w) = a\mu(v, w) + b\mu(v', w) = av\otimes w + bv'\otimes w,$$

and similarly

$$v\otimes (aw + bw') = av\otimes w + bv\otimes w'.$$

1.4 Proposition.

If F : V × W → X is any bilinear map, then F can be written uniquely as composition

$$V \times W \xrightarrow{\mu} V \otimes W \xrightarrow{\underline{F}} X,$$

whereμ is the multiplication map, and $\underline{F}$ is linear map.

Let's first note we already know that$\underline{F}$: V ⊗ W has to be on pure tensors:

$$\underline{F}(v \otimes w) = \underline{F}(\phi(v, w)) = F(v, w).$$

Since a general vector in V ⊗ W is a sum of some pure tensors, and we want $\underline{F}$ to be linear, it is determined by its value on pure tensors (extending to all of V ⊗ W by linearity). The proposition guarantees that the result $\underline{F}$ is indeed a linear map.

1.5 Corollary.

Let L(V×W, X) denote the space of bilinear maps from V×W into X. Then there is an isomorphism L(V × W, X) ≅ L(V ⊗ W, X).

where the right hand side denotes the space of linear maps out of the tensor product.

The above discussion gives us two ways of constructing linear maps out of the tensor product of vector spaces:

(a) Just define the values of the map on a basis $\{v_i \otimes w_j\}$, and extend by linearity; or
(b) Define the T on pure tensors, of the form $v \otimes w$, and extend by linearity to all of $V \otimes W$.

By Proposition 1.4, if the map

$$\hat{T} : V \times W \to X$$

$$(v, w) \mapsto T(v \otimes w)$$

is bilinear, then T is linear.

Let V, W be finite dimensional vector spaces over F. Let V*= L(V, F) be the dual space of V . Then $V^* \otimes W = L(V, W)$, with multiplication defined as $f \otimes w \in L(V, W)$ is the linear transformation $(f \otimes w)(v) = f(v)w$, for $f \in V^*, w \in W$.

Note: If V and Ware both infinite dimensional then $V^* \otimes W$ is a subspace of L(V, W) but not equal to it.If $V = F^n_{col}$ and $W = F^m_{col}$, then $V^* = F^n_{row}$ and $V \otimes W$ is identified with $M_{m\times n}(F)$, which is in turn identified with L(V, W).

1.6 Definition.

If A is an m × n matrix and B is a p × q matrix, then the **Matrix tensor product**$A \otimes B$ is the mp × nq block matrix:

$$A \otimes B = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

The following properties hold for the matrix tensor product:

(i) $A \otimes (B + C) = A \otimes B + A \otimes C$

(ii) $(A + B) \otimes C = A \otimes C + B \otimes C$

(iii) $(kA) \otimes B = A \otimes (kB) = k(A \otimes B)$

(iv) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$

whereA, B and C are matrices and k is a scalar.

If A, B, C and D are matrices of such size that one can form the matrix products AC and BD, then $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

If S and T are square matrices of order m and n respectively, we define the linear mapping $S \otimes T : M_{m\times n}(F) \to M_{m\times n}(F)$ by linear extension of

$(S \otimes T)(e_i \otimes f_j) = (Se_i \otimes Tf_j)$, where $\{e_i\}$ and $\{f_j\}$ are standard bases, i = 1, 2, …, m, j = 1, 2,…, n. Then the linear mapping $S \otimes T$ is the tensor product of the matrices S and T.

If $x \in F^m_{col}$ and $y \in F^n_{col}$, then $(S \otimes T)(x \otimes y) = (Sx) \otimes (Ty) = Sx(Ty)^T$.

2. Applications

2.1 In classical mechanics

The inertia matrix is often described as the inertia tensor, which consists of the same moments of inertia and products of inertia about the three coordinate axes. The inertia tensor is constructed from the nine component tensors, (the symbol $\otimes$ is the tensor product)

$$e_i \otimes e_j, \quad i, j = 1, 2, 3,$$

where $e_i, i = 1, 2, 3$ are the three orthogonal unit vectors defining the inertial frame in which the body moves. Using this basis the inertia tensor is given by

$$I = \sum_{i=1}^{3}\sum_{j=1}^{3} I_{ij} e_i \otimes e_j$$

This tensor is of degree two because the component tensors are each constructed from two basis vectors.

For a rigid system of particles $P_k, k = 1, ..., N$ each of mass m_k with position coordinates $r_k = (x_k, y_k, z_k)$, the inertia tensor is given by

$$I = \sum_{k=1}^{N} m_k \left((r_k \cdot r_k) E - r_k \otimes r_k \right), \text{where } E = e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3.$$

In this case, the components of the inertia tensor are given by

$$I_{11} = I_{xx} = \sum_{k=1}^{N} m_k \left(y_k^2 + z_k^2 \right),\ I_{22} = I_{yy} = \sum_{k=1}^{N} m_k \left(x_k^2 + z_k^2 \right), I_{33} = I_{zz} = \sum_{k=1}^{N} m_k \left(x_k^2 + y_k^2 \right)$$

$$I_{12} = I_{21} = I_{xy} = -\sum_{k=1}^{N} m_k x_k y_k, I_{13} = I_{31} = I_{xz} = -\sum_{k=1}^{N} m_k x_k z_k, I_{23} = I_{32} = I_{yz}$$

$$= -\sum_{k=1}^{N} m_k y_k z_k.$$

2.2 In quantum mechanics

The quantum mechanical operator associated with spin-$\frac{1}{2}$ observables is $\hat{\mathbf{S}} = \frac{\hbar}{2}\boldsymbol{\sigma}$ where in Cartesian components:

$$S_x = \frac{\hbar}{2}\sigma_x , S_y = \frac{\hbar}{2}\sigma_y , S_z = \frac{\hbar}{2}\sigma_z .$$

For the special case of spin-$\frac{1}{2}$ particles, σ_x, σ_y and σ_z are the three Pauli matrices, given by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We can construct higher irreducible representations by taking matrix tensor products of the representation $\hat{S}$ with itself repeatedly.

The resulting spin 1 matrices and eigenvalues in the z-basis are

$$S_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, S_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \text{ and } S_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

2.3 In differential equations

The stack operator maps an $n \times m$ matrix into an $nm \times 1$ vector. The stack of the $n \times m$ matrix A, denoted A^S, is the vector formed by stacking the columns of A into an $nm \times 1$ vector.

The Stack Operator has the following properties.

(i) If $v \in R^{n\times 1}$, a vector, then $v^S = v$.

(ii) If $A \in R^{m\times n}$, a matrix, and $v \in R^{n\times 1}$, a vector, then the matrix product $(Av)^S = Av$.

Stack of a matrix multiplication, when dimensions are appropriate for the product ABC to be welldefined, is

$$(ABC)^S = (C^T \otimes A)B^S,$$

and

$$Ax = IAx = (x^T \otimes I)A^S.$$

For any Hurwitz $A \in R^{n\times n}$and any positive-definite symmetricQ $\in R^{n\times n}$there exists a unique positive-definite symmetric $P \in R^{n\times n}$satisfying the linear Lyapunovequation

– Q = AT P + PA.

The matrix P can be computed as

$$-Q = A^T P + PA = A^T PI + IPA$$

$$-Q^S = (I \otimes A^T)P^S + (A^T \otimes I)P^S$$

$$-Q^S = (I \otimes A^T + A^T \otimes I)P^S$$

thus$P^S$is given by

$$P^S = -(I \otimes A^T + A^T \otimes I)^{-1}Q^S$$

The derivative of vector-values and matrix-valued functions of vectors or matrices such as $y = A(b)c$,

where $y \in R^n$, $b \in R^m$, $c \in R^p$, and $A : R^m \to R^{n\times p}$, can be written as

$$\dot{y} = \frac{d}{dt}[A(b)c]$$

$$= A(b)\dot{c} + \frac{d}{dt}[A(b)]c$$

$$= A(b)\dot{c} + I\frac{d}{dt}[A(b)]c$$

$$= A(b)\dot{c} + [I\frac{d}{dt}[A(b)]c]^S$$

$$= A(b)\dot{c} + (c^T\otimes I)\frac{d}{dt}[A(b)^S]$$

$$= A(b)\dot{c} + (c^T\otimes I)D_b[A(b)^S]\dot{b}$$

where$D_b[A(b)^S]$ is the usual matrix-valued jacobian operator.

2.4 In graph theory

We consider two directed graphs $G_1=(\{v_1,v_2,v_3\},\{v_1v_2,\ v_1v_3,\ v_3v_2\})$ and $G_2 = (\{u_1,u_2\}\{u_1u_2\})$.

We can make two adjacency matrices for G_1 and G_2:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then $A\otimes B = \begin{pmatrix} 0&0&0&1&0&1\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&0 \end{pmatrix}$ is an adjacency matrix for directed graph

$G_1\otimes G_2=(P, Q)$, where

$P = \{(v_1,u_1), (v_1, u_2), (v_2, u_1), (v_2, u_2), (v_3, u_1), (v_3, u_2)\}$ and

$Q = \{(v_1,u_1)(v_2, u_2), (v_1,u_1)(v_3, u_2),(v_3, u_1)(v_2, u_2)\}$.

2.5 Smoothing an image

We want to smooth the image x with the one-dimensional filters S and T given by $y_n = (Sx)_n = (Tx)_n = (x_{n-1} + 2x_n + x_{n+1})/4$. Set m = 5 and n = 7. Then

$$S = \frac{1}{4}\begin{pmatrix} 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 \end{pmatrix} \text{ and } T = \frac{1}{4}\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Now we compute $(S \otimes T)(e_2 \otimes f_3)$.

$(S \otimes T)(e_2 \otimes f_3) = (Se_2 \otimes Tf_3) = Se_2(Tf_3)^T = (col_2S)(col_3T)^T$.

Since $Se_2 = = col_2S$ and $Tf_3 = col_3T$

$$\text{we find that} Se_2(Tf_3)^T = \frac{1}{16}\begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 4 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (col_2S)(col_3T)^T.$$

References

[1] Godsil, C., andRoyle, G. F., “Algebraic Graph Theory”, Springer-Verlag, New York.

[2] Jain, A. K., (1989), "Fundamentals of Digital Image Processing" Prentice Hall.

[3] Lang, S., (1987), “Linear Algebra”, Springer-Verlag New York, Inc.

[4] Schafer, H. N., and Schmidt, J. P., (2017), "Tensor Analysis and Elementary Differential Geometry for Physicists and Engineers", 2nd edition, Springer-Verlag, Berlin.