A conceptual proof of the identity $\operatorname{vec}(ABC) = (A^t \otimes C)^t \operatorname{vec}(B)$

Bart Michels

April 10, 2018

We work over a commutative ring (with unity) R. Let $M_{p,q}$ denote the $p \times q$ matrix space over R, which we identify with the space of linear maps $R^q \to R^p$ by interpreting them as matrices of linear maps in the standard bases. Let E_{ij} be the standard basis matrix with a 1 on position (i, j) and 0 elsewhere. The size of $E_{i,j}$ depends on the context. We order them using the lexicographical order, that is, first by row, then by column.

Definition 1. The **(row) vectorization** $\operatorname{vec}(A)$ of a matrix A is the column vector obtained by stacking its (transposed) rows. It is its coordinate vector in the standard basis $(E_{i,j})$.

We study how matrix products behave with respect to vectorization.

Remark. Usually the name 'vectorization' is used for what one could call the **column vectorization** and the (E_{ij}) are ordered colexicographically. This turns out to be unnatural in this context; see Lemma 4. They are related by $\operatorname{vec}^c(A) = \operatorname{vec}^r(A^t)$.

Definition 2. For free R-modules P, Q with bases (e_i) and (f_j), the **tensored basis** of $P \otimes Q$ is the one consisting of $e_i \otimes f_j$, ordered lexicographically, that is, first by i then by j.

Definition 3. For $A \in M_{p,q}$ and $B \in M_{r,s}$, their **Kronecker product** is the $pr \times qs$ matrix of the tensored map $A \otimes B : R^q \otimes R^s \to R^p \otimes R^r$ in the standard bases. Explicitly, it is given by the block matrix

$$\begin{pmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{pmatrix}$$

Remark. With this definition we immediately have identities like $AB \otimes CD = (A \otimes C)(B \otimes D)$ whenever the LHS is defined. We won't need those.

The vectorization and the Kronecker product give the impression that they are related. Indeed, the heart of the matter is that matrices are identified with bilinear maps, which are identified with elements of the dual of a tensor product, and thus with elements of the tensor product:

Lemma 4. Let $A \in M_{p,q}$ be considered as a bilinear map on $R^p \times R^q$, that is, an element of the dual $(R^p \otimes R^q)^*$, which we identify via the choice of a basis with $R^p \otimes R^q$. Then $\operatorname{vec}(A)$ is the coordinate vector of A as an element of $R^p \otimes R^q$.

Proof. We only have to check that E_{ij} goes to $e_i \otimes f_j$ with those identifications, since the ordering of the bases is the same. Indeed, $E_{ij} = e_i f_j^t$ is the bilinear map that sends (e_i, f_j) to 1 and all other (e_i, f_j) to 0. Thus as an element of the dual $(R^p \otimes R^q)^*$ it sends $e_i \otimes f_j$ to 1 and all other tensor basis elements to 0. By definition of the identification of a module with its dual via a basis, this linear form is identified with $e_i \otimes f_j$. □

Theorem 5. Let A, B, C be matrices such that ABC is defined. Then $\operatorname{vec}(ABC) = (A \otimes C^t) \operatorname{vec}(B)$.

The proof will explain why the dimensions match for the statement to make sense.

Proof. We interpret ABC and B as bilinear maps, and A and C as linear maps. On the level of bilinear maps, we have a commutative diagram:

$$\begin{tikzcd} R^s \times R^t \arrow[r, "{(A^t, C)}"] \arrow[rr, bend right, "ABC"'] & R^p \times R^q \arrow[r, "B"] & R \end{tikzcd}$$

On the level of duals of tensor products, we have:

$$\begin{tikzcd} R^s \otimes R^t \arrow[r, "A^t \otimes C"] \arrow[rr, bend right, "B \circ (A^t \otimes C) = ABC"'] & R^p \otimes R^q \arrow[r, "B"] & R \end{tikzcd}$$

where the equality $B \circ (A^t \otimes C) = ABC$ follows from the universal property of $R^s \otimes R^t$. By definition of the duality functor, precomposition with $A^t \otimes C$ is the dual of $A^t \otimes C$ as a linear map. By identifying a free module with its dual, we revert this functor on objects and take the transpose of linear maps (well-defined when fixing a basis), and so on the level of tensor products, we have:

$$\begin{aligned} R^s \otimes R^t &\xleftarrow{(A^t \otimes C)^t} R^p \otimes R^q \\ ABC &\longleftarrow\!\shortmid\; B \end{aligned}$$

So that by Lemma 4 and the definition of Kronecker product:

$$\mathrm{vec}(ABC) = (A^t \otimes C)^t \,\mathrm{vec}(B)$$

It remains to check that $(A^t \otimes C)^t = A \otimes C^t$, which is done in the lemma below. □

Lemma 6. We have $(A \otimes B)^t = A^t \otimes B^t$ whenever the LHS is defined.

Proof. meh

□