The Kronecker Product

Definition (Kronecker product)

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{p \times r}$. The *Kronecker product* between $\mathbf{A}$ and $\mathbf{B}$ is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix} \in \mathbb{R}^{mp \times nr}.$$

Note that, in general, $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$ (i.e., the Kronecker product is not commutative).

The Kronecker Product (Some Properties)

The Kronecker product has many nice properties, for example:

(K1) Linearity: $\left(\sum_k c_k \mathbf{A}_k\right) \otimes \mathbf{B} = \sum_k c_k (\mathbf{A}_k \otimes \mathbf{B})$

(K2) Associativity: $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$

(K3) Mixed product: $(\mathbf{AC}) \otimes (\mathbf{BD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D})$

Proving properties K2 and K3 is a nice exercise in block matrix multiplication!

Vectorisation

Definition (Vectorisation)

Let $\mathbf{A} = \begin{pmatrix}\mathbf{A}_1 & \cdots & \mathbf{A}_n\end{pmatrix} \in \mathbb{R}^{m\times n}$, where each $\mathbf{A}_k \in \mathbb{R}^m$. The *vectorisation* of $\mathbf{A}$ is defined as

$$\operatorname{vec}\mathbf{A} = \begin{pmatrix}\mathbf{A}_1 \\ \vdots \\ \mathbf{A}_n\end{pmatrix} \in \mathbb{R}^{mn}.$$

(V1) Linearity: $\operatorname{vec}\left(\sum_k c_k \mathbf{A}_k\right) = \sum_k c_k \operatorname{vec}\mathbf{A}_k$

(V2) Outer product: For $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$, we have

$$\operatorname{vec}(\mathbf{u}\mathbf{v}^\top) = \operatorname{vec}\begin{pmatrix} v_1\mathbf{u} & \cdots & v_n\mathbf{u}\end{pmatrix} = \begin{pmatrix} v_1\mathbf{u} \\ \vdots \\ v_n\mathbf{u}\end{pmatrix} = \mathbf{v} \otimes \mathbf{u}.$$

Vectorisation of Matrix Product

Theorem

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{X} \in \mathbb{R}^{n \times p}$, *and* $\mathbf{C} \in \mathbb{R}^{p \times r}$. *Then*

$$\operatorname{vec}(\mathbf{AXC}) = \left(\mathbf{C}^\top \otimes \mathbf{A}\right) \operatorname{vec} \mathbf{X}.$$

This theorem is *very* useful for solving linear matrix equations in a systematic way! If we want to solve for $\mathbf{X}$ such that $\mathbf{AXC} = \mathbf{Y}$ (where $\mathbf{A}$ and $\mathbf{C}$ may not even be square), the theorem transforms the problem into solving a standard form system

$$\tilde{\mathbf{A}}\tilde{\mathbf{X}} = \tilde{\mathbf{Y}},$$

with $\tilde{\mathbf{A}} = \mathbf{C}^\top \otimes \mathbf{A}$, $\quad \tilde{\mathbf{X}} = \operatorname{vec} \mathbf{X}$, and $\tilde{\mathbf{Y}} = \operatorname{vec} \mathbf{Y}$.

Vectorisation of Matrix Product, Proof

Proof.

We start with two preliminary observations.

First, let $\mathbf{I} = \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_p \end{pmatrix}$ be the $p \times p$ identity matrix, and note that

$$\mathbf{X} = \mathbf{X}\mathbf{I}^\top = \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_p \end{pmatrix} \begin{pmatrix} \mathbf{e}_1^\top \\ \vdots \\ \mathbf{e}_p^\top \end{pmatrix} = \sum_{k=1}^{p} \mathbf{x}_k \mathbf{e}_k^\top .$$

Secondly, note that

$$\operatorname{vec} \mathbf{X} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_p \end{pmatrix} = \sum_{k=1}^{p} (\mathbf{e}_k \otimes \mathbf{x}_k).$$

Vectorisation of Matrix Product, Proof (contd.)

Proof (contd.)

Using the first observation to rewrite $\mathbf{X}$, we have

$$\operatorname{vec}(\mathbf{AXC}) = \operatorname{vec}\left(\mathbf{A}\left(\sum_{k=1}^{p} \mathbf{x}_k \mathbf{e}_k^\top\right)\mathbf{C}\right) = \operatorname{vec}\left(\sum_{k=1}^{p} \mathbf{A}\mathbf{x}_k \mathbf{e}_k^\top \mathbf{C}\right) = \Big/ \text{use (V1)} \Big/ =$$

$$= \sum_{k=1}^{p} \operatorname{vec}\left(\mathbf{A}\mathbf{x}_k \mathbf{e}_k^\top \mathbf{C}\right) = \sum_{k=1}^{p} \operatorname{vec}\left(\mathbf{A}\mathbf{x}_k (\mathbf{C}^\top \mathbf{e}_k)^\top\right) = \Big/ \text{use (V2)} \Big/ =$$

$$= \sum_{k=1}^{p} (\mathbf{C}^\top \mathbf{e}_k) \otimes (\mathbf{A}\mathbf{x}_k) = \Big/ \text{use (K3)} \Big/ = \sum_{k=1}^{p} (\mathbf{C}^\top \otimes \mathbf{A})(\mathbf{e}_k \otimes \mathbf{x}_k).$$

Bringing $\mathbf{C}^\top \otimes \mathbf{A}$ outside the sum, and using observation two, proves the result. $\square$