A Note on the vec Operator Applied to Unbalanced Block-Structured Matrices

Caswell, H.; van Daalen, S.F.

Published in:
Journal of applied mathematics

DOI:
10.1155/2016/4590817

Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2016, Article ID 4590817, 3 pages
http://dx.doi.org/10.1155/2016/4590817

Research Article

A Note on the vec Operator Applied to Unbalanced Block-Structured Matrices

Hal Caswell and Silke F. van Daalen

Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, Netherlands

Correspondence should be addressed to Hal Caswell; h.caswell@uva.nl

Received 25 July 2016; Accepted 3 October 2016

Academic Editor: Frédéric Guichard

The vec operator transforms a matrix to a column vector by stacking each column on top of the next. It is useful to write the vec of a block-structured matrix in terms of the vec operator applied to each of its component blocks. We derive a simple formula for doing so, which applies regardless of whether the blocks are of the same or of different sizes.

1. Introduction

The vec operator, applied to a $r \times c$ matrix $\mathbf{X}$, produces a $rc \times 1$ column vector, denoted $\mathrm{vec}(\mathbf{X})$ by stacking each column of $\mathbf{X}$ on top of the following column [1]. Here, we consider the result of applying the vec operator to block-structured matrices, including the case in which the blocks differ in size. Such a matrix is called unbalanced [2]. Previous studies of the vec operator and Kronecker product applied to block-structured matrices [2, 3] have not addressed this problem.

In many applications, block-structured matrices arise because the blocks represent different states or processes. In general, these blocks will be of different sizes and may depend on different parameters. If the vec operator is applied to such a matrix, it may be helpful to write the result in terms of the vec of each of the component blocks. This calculation arises, *inter alia*, in applications of matrix calculus [4] in demography and ecology, including nonlinear matrix population models [5] and finite-state Markov chains [6, 7]. In such models (we give an example below), the outcome is often a vector-valued function of the same matrix, and the matrix has an inherent block structure.

Our goal is to write the vec of the unbalanced block-structured matrix as a linear combination of the vec operator applied to each of the component blocks. Although the solution is simple, it is widely useful, so we present it here.

2. Results

If the matrix $\mathbf{X}$ contains n blocks, we write it as the sum of n matrices, each containing one of the blocks surrounded by zero matrices, as in

$$\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \multicolumn{2}{c}{\mathbf{C}} \end{array}\right) \tag{1}$$

$$= \left(\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \multicolumn{2}{c}{\mathbf{0}} \end{array}\right) + \left(\begin{array}{c|c} \mathbf{0} & \mathbf{B} \\ \hline \multicolumn{2}{c}{\mathbf{0}} \end{array}\right) + \left(\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \multicolumn{2}{c}{\mathbf{C}} \end{array}\right), \tag{2}$$

where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ and the corresponding zero matrices may be of different (but compatible) sizes. The vec of $\mathbf{X}$ is the sum of the vec operator applied to each of the component matrices in (2). A generic member of this set of component matrices can be written as in the following result.

Theorem 1. *Let $\mathbf{X}$ be an $r \times c$ block-structured matrix, with the dimensions of the blocks indicated as subscripts, written*

$$\mathbf{X}_{r\times c} = \left(\begin{array}{c|c|c} \mathbf{0}_{r_1\times c_1} & \mathbf{0}_{r_1\times c_2} & \mathbf{0}_{r_1\times c_3} \\ \hline \mathbf{0}_{r_2\times c_1} & \mathbf{A}_{r_2\times c_2} & \mathbf{0}_{r_2\times c_3} \\ \hline \mathbf{0}_{r_3\times c_1} & \mathbf{0}_{r_3\times c_2} & \mathbf{0}_{r_3\times c_3} \end{array}\right), \tag{3}$$

where $r = r_1 + r_2 + r_3$ and $c = c_1 + c_2 + c_3$ and r_1, r_3, c_1, c_3, or any combination, may be zero. Then

$$\mathrm{vec}\,(\mathbf{X}) = \left(\mathbf{Q}^{\top} \otimes \mathbf{P}\right) \mathrm{vec}\,\mathbf{A}, \tag{4}$$

where

$$\mathbf{P} = \begin{pmatrix} \mathbf{0}_{r_1\times r_2} \\ \mathbf{I}_{r_2\times r_2} \\ \mathbf{0}_{r_3\times r_2} \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{0}_{c_2\times c_1} & \mathbf{I}_{c_2\times c_2} & \mathbf{0}_{c_2\times c_3} \end{pmatrix}. \tag{5}$$

Proof. To convert $\mathbf{A}$ to $\mathbf{X}$ requires the addition of r_1 rows of zeros above, r_3 rows of zeros below, c_1 columns of zeros to the left, and c_3 columns of zeros to the right of $\mathbf{A}$. This is accomplished by multiplying $\mathbf{A}$ on the left by $\mathbf{P}$ and on the right by $\mathbf{Q}$, so that

$$\mathbf{X} = \mathbf{PAQ}. \tag{6}$$

Applying the vec operator to (6), using a well known result of Roth [8], yields (4). □

Remark 2. We said it was simple.

3. Applications

Here are several examples of interest, to demonstrate the formulation of the block-structured matrices and the result of applying the vec operator.

(1) The transition matrix for a finite-state absorbing Markov chain with τ transient states and α absorbing states can be written as a block-structured matrix. Numbering the states so that the transient states precede the absorbing states yields a canonical form for the (column-stochastic) transition matrix (e.g., [9]),

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U}_{\tau\times\tau} & \mathbf{0}_{\tau\times\alpha} \\ \hline \mathbf{M}_{\alpha\times\tau} & \mathbf{I}_{\alpha\times\alpha} \end{array}\right). \tag{7}$$

The matrix $\mathbf{U}$ describes transitions among the transient states and $\mathbf{M}$ describes transitions from transient states to absorbing states. Suppose that $\boldsymbol{\xi}$ is a vector-valued $(s \times 1)$ differentiable function of $\mathbf{P}$ and that $\mathbf{U}$ and $\mathbf{M}$ are differentiable functions of a vector $(p \times 1)$ $\boldsymbol{\theta}$ of parameters. In demographic and ecological applications, $\mathbf{U}$ might describe transitions and survival among life cycle stages of some organism, and $\mathbf{M}$ might describe transitions to different causes of death (e.g., [7, 10]).

Following [4], the derivative of $\boldsymbol{\xi}$ with respect to $\boldsymbol{\theta}$ is the $s \times p$ matrix

$$\frac{d\boldsymbol{\xi}}{d\boldsymbol{\theta}^\top} = \frac{d\boldsymbol{\xi}}{d\,\mathrm{vec}^\top\mathbf{P}} \frac{d\,\mathrm{vec}\,\mathbf{P}}{d\boldsymbol{\theta}^\top}. \tag{8}$$

To obtain $d\,\mathrm{vec}\,\mathbf{P}$ we must apply the vec operator to the block-structured matrix $d\mathbf{P}$. Applying the results (5) gives

$$\begin{aligned} d\mathbf{P} = & \begin{pmatrix} \mathbf{I}_{\tau\times\tau} \\ \mathbf{0}_{\alpha\times\tau} \end{pmatrix} d\mathbf{U} \begin{pmatrix} \mathbf{I}_{\tau\times\tau} & \mathbf{0}_{\tau\times\alpha} \end{pmatrix} \\ & + \begin{pmatrix} \mathbf{0}_{\tau\times\alpha} \\ \mathbf{I}_{\alpha\times\alpha} \end{pmatrix} d\mathbf{M} \begin{pmatrix} \mathbf{I}_{\tau\times\tau} & \mathbf{0}_{\tau\times\alpha} \end{pmatrix} \end{aligned} \tag{9}$$

(noting that $d\mathbf{I}$ and $d\mathbf{0}$ are both zero). Applying the vec operator and the chain rule gives

$$\begin{aligned} \frac{d\,\mathrm{vec}\,\mathbf{P}}{d\boldsymbol{\theta}^\top} = & \left[\begin{pmatrix} \mathbf{I}_{\tau\times\tau} \\ \mathbf{0}_{\alpha\times\tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{I}_{\tau\times\tau} \\ \mathbf{0}_{\alpha\times\tau} \end{pmatrix}\right] \frac{d\,\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^\top} \\ & + \left[\begin{pmatrix} \mathbf{I}_{\tau\times\tau} \\ \mathbf{0}_{\alpha\times\tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{0}_{\tau\times\alpha} \\ \mathbf{I}_{\alpha\times\alpha} \end{pmatrix}\right] \frac{d\,\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^\top}. \end{aligned} \tag{10}$$

In applications, it is likely that parameters of interest are defined in terms of their effects on $\mathbf{U}$ and $\mathbf{M}$; (10) makes it possible to incorporate that dependence easily into the necessary derivative of the block-structured matrix $\mathbf{P}$.

We note that the intensity matrix of a continuous-time absorbing Markov chain also has a block structure (e.g., [9, Chap. 8]); applications of matrix calculus to such models [6] will benefit from the results presented here.

(2) The transition matrix of an absorbing Markov chain is a special case of the canonical form of a reducible nonnegative matrix [11]. If $\mathbf{X}$ is a reducible matrix, it can be written

$$\mathbf{X} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{s1} & \mathbf{A}_{s2} & \cdots & \mathbf{A}_{ss} \end{pmatrix}. \tag{11}$$

Each of the diagonal blocks $\mathbf{A}_{ii}$ is square and irreducible (or a 1×1 zero matrix). The division into diagonal blocks corresponds to a division of the vector space upon which $\mathbf{X}$ operates into invariant subspaces.

(3) A balanced block-structured matrix, in which all blocks are the same size, is a special case of an unbalanced matrix. Theorem 1 provides a simple result for the vec of such a matrix. Consider the $pm \times qn$ matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1q} \\ \vdots & & \vdots \\ \mathbf{A}_{p1} & \cdots & \mathbf{A}_{pq} \end{pmatrix}, \tag{12}$$

where $\mathbf{A}_{ij}$ are each of dimension $m \times n$. This case is considered by [3].

From (5)-(6) we have

$$\mathbf{X} = \sum_{i=1}^{p}\sum_{j=1}^{q} \mathbf{P}_i \mathbf{A}_{ij} \mathbf{Q}_j, \tag{13}$$

where

$$\begin{aligned} \mathbf{P}_i &= \begin{pmatrix} \mathbf{0}_{(i-1)m\times m} \\ \mathbf{I}_{m\times m} \\ \mathbf{0}_{(p-i)m\times m} \end{pmatrix}, \quad i = 1, \ldots, p, \\ \mathbf{Q}_j &= \begin{pmatrix} \mathbf{0}_{n\times(j-1)n} & \mathbf{I}_{n\times n} & \mathbf{0}_{n\times(q-j)n} \end{pmatrix}, \quad j = 1, \ldots, q. \end{aligned} \tag{14}$$

The vec of $\mathbf{X}$ is

$$\mathrm{vec}\,\mathbf{X} = \sum_{i=1}^{p}\sum_{j=1}^{q} \left(\mathbf{Q}_j^\top \otimes \mathbf{P}_i\right) \mathrm{vec}\,\mathbf{A}_{ij}. \tag{15}$$

4. Conclusions

The vec operator, by transforming a matrix into a vector, is useful in many applications [1]. When the matrix is block structured and the blocks represent various processes involved in the application, it is convenient to be able to express the vec of the matrix as a linear combination of the vec operator applied to the component blocks. We have shown how to do so and described a few examples, but these do not exhaust the potential uses of the result.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by ERC Advanced Grant 322989, NSF Grant DEB-1257545, and NWO-ALW Project ALWOP.2015.100.

References

[1] H. V. Henderson and S. R. Searle, "The vec-permutation matrix, the vec operator and Kronecker products: a review," *Linear and Multilinear Algebra*, vol. 9, no. 4, pp. 271–288, 1981.

[2] R. H. Koning, H. Neudecker, and T. Wansbeek, "Block Kronecker products and the vecb operator," *Linear Algebra and Its Applications*, vol. 149, pp. 165–184, 1991.

[3] I. Ojeda, "Kronecker square roots and the block vec matrix," *The American Mathematical Monthly*, vol. 122, no. 1, pp. 60–64, 2015.

[4] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, New York, NY, USA, 1988.

[5] H. Caswell, "Sensitivity and elasticity of density-dependent population models," *Journal of Difference Equations and Applications*, vol. 15, no. 4, pp. 349–369, 2009.

[6] H. Caswell, "Perturbation analysis of continuous-time absorbing Markov chains," *Numerical Linear Algebra with Applications*, vol. 18, no. 6, pp. 901–917, 2011.

[7] H. Caswell, "Sensitivity analysis of discrete Markov chains via matrix calculus," *Linear Algebra and Its Applications*, vol. 438, no. 4, pp. 1727–1745, 2013.

[8] W. E. Roth, "On direct product matrices," *Bulletin of the American Mathematical Society*, vol. 40, no. 6, pp. 461–468, 1934.

[9] M. Iosifescu, *Finite Markov Processes and Their Applications*, Wiley-Interscience, New York, NY, USA, 1980.

[10] H. Caswell, "Stage, age and individual stochasticity in demography," *Oikos*, vol. 118, no. 12, pp. 1763–1782, 2009.

[11] F. R. Gantmacher, *Matrix Theory, Volume 2*, Chelsea, New York, NY, USA, 1959.