The American Mathematical Monthly

ISSN: 0002-9890 (Print) 1930-0972 (Online) Journal homepage: http://www.tandfonline.com/loi/uamm20

Rank My Update, Please

Carl D. Meyer

To cite this article: Carl D. Meyer (2018) Rank My Update, Please, The American Mathematical Monthly, 125:1, 61-64, DOI: 10.1080/00029890.2017.1389199

To link to this article: https://doi.org/10.1080/00029890.2017.1389199

Published online: 21 Dec 2017.

Rank My Update, Please

Carl D. Meyer

Abstract. Updating a given a matrix $\mathbf{A}_{m\times n}$ by a rank-one matrix $\mathbf{B} = \mathbf{c}\mathbf{d}^T$, where $\mathbf{c}$ and $\mathbf{d}$ are appropriately sized column vectors, is a common practice throughout all applied areas of mathematics, science, and engineering. Because rank is often tied to the number of degrees of freedom or the level of independence in underlying models or data, it can be imperative to know exactly how the update term affects rank. While it is well known that a rank-one update can only increase or decrease rank by at most one, there is not a widely known formula for exactly how this occurs. This note presents an expression in simply stated terms for the exact rank of a rank-one updated matrix.

1. INTRODUCTION. "Rank" is a fundamental concept in both the theory and applications of linear algebra, and elementary properties of rank are always focal points in any undergraduate course. Among the first things that students learn when considering matrices is that the rank of a product is generally not equal to the product of the ranks, and they are usually directed to the next best result, which is that the rank of a product does not exceed the rank of any factor—i.e.,

$$rank\,(\mathbf{AB}) \le \min\{rank\,(\mathbf{A}),\ rank\,(\mathbf{B})\}. \tag{1}$$

The fact that there is an exact formula for the rank of a product is frequently hidden from beginning students until they move to a higher level, and even then it is not common in popular texts to expose them to the useful formula from [**2**, p. 210] that states

$$rank\,(\mathbf{AB}) = rank\,(\mathbf{B}) - \dim N\,(\mathbf{A}) \cap R\,(\mathbf{A}), \tag{2}$$

where $R(\star)$ and $N(\star)$, respectively, denote range and nullspace. (Note that the inequality in (1) is an easy consequence of (2)).

And there is a similar theme for sums. Students learn that the rank of a sum is generally not the sum of the ranks—i.e., rank is not linear but rather sub-linear in the sense that

$$rank\,(\mathbf{A}+\mathbf{B}) \le rank\,(\mathbf{A}) + rank\,(\mathbf{B}). \tag{3}$$

However, unlike the situation for a product, there seems to be no useful elementary formula for the exact rank of a sum that is analogous to formula (2) for a product. Consequently, attention is again directed to the next best possibilities, which means that we either revert to establishing useful bounds for the rank of a sum, or we limit attention to determining expressions for the exact rank of special but frequently occurring sums. The point of this note is the latter, but first some preliminaries concerning the former are mentioned.

doi.org/10.1080/00029890.2017.1389199
MSC: Primary 15-01

2. UPDATES. In applications, sums $\mathbf{A}+\mathbf{B}$ often arise when $\mathbf{B}$ contains updates or perturbations to prescribed data or information stored in $\mathbf{A}$. Knowing $rank(\mathbf{A}+\mathbf{B})$ is these situations can be important because rank often represents the number of degrees of freedom or the levels of independence in an underlying physical model or in data contained in $\mathbf{A}$, so the degree to which $\mathbf{B}$ changes $rank(\mathbf{A})$ can have significant implications. In these applications, the issues are generally not the same as gauging the effects of small perturbations because the update term $\mathbf{B}$ need not be small in *magnitude* relative to $\|\mathbf{A}\|$, but rather it is the *rank* of the update that is often small relative to that of $\mathbf{A}$.

Bounding the effect of updates on rank is not difficult. It is a relatively straightforward exercise [**2**, p. 208] to use the inequality in (3) to show that

$$|rank\,(\mathbf{A}) - rank\,(\mathbf{B})| \leq rank\,(\mathbf{A}-\mathbf{B}),$$

which in turn leads to the conclusion that if $rank(\mathbf{A}) = r$ and $rank(\mathbf{B}) = k$, then

$$r - k \leq rank\,(\mathbf{A}+\mathbf{B}) \leq r + k \tag{4}$$

so that the effect of a rank-k update to $\mathbf{A}$ is to either increase or decrease its rank by at most k.

3. RANK-ONE UPDATES. An obvious consequence of (4) is that the effect of an update of minimal rank—i.e., a rank-one update—is to increase or decrease the rank of $\mathbf{A}_{m\times n}$ by at most one, and since rank-one matrices are of the form[1] $\mathbf{B}_{m\times n} = \mathbf{c}\mathbf{d}^T$ in which $\mathbf{c}_{m\times 1} \neq \mathbf{0}$ and $\mathbf{d}_{n\times 1} \neq \mathbf{0}$, inequality (4) reduces to

$$r - 1 \leq rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T) \leq r + 1. \tag{5}$$

Applications involving rank-one updates are ubiquitous throughout science, engineering, and applied mathematics—e.g, a Google search on "rank-one update" can return thousands of results.

The most familiar updating results are the Sherman–Morrison theorem and its generalization known as the Sherman–Morrison–Woodbury theorem [**2**, p. 124]. When $rank(\mathbf{A}_{\mathbf{n}\times\mathbf{n}}) = n$, the Sherman–Morrison theorem states that $rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T) = n$ if and only if $\alpha = 1 + \mathbf{d}^T\mathbf{A}\mathbf{c} \neq 0$. Furthermore,

$$(\mathbf{A}+\mathbf{c}\mathbf{d}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^T\mathbf{A}^{-1}}{\alpha}.$$

But when $\mathbf{A}$ is square and singular or rectangular and rank deficient, the simple elegance of the Sherman–Morrison theorem is lost in the weeds [**1**, pp. 47–48, 116], and just providing a simply-stated formula for $rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T)$ is challenging. As previously mentioned, the bound in (5) will not suffice when it is important to know if (or how) an update will increase, maintain, or decrease rank to gauge possible changes relevant to underlying models or data, and this is the motivation for this note. The goal is to use elementary terms to exhibit explicit expressions for $rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T)$ that seem not to be widely known, if at all.

4. SVD IS THE KEY. A singular value decomposition (or SVD) of a matrix is capable of providing an inordinate number of facts and properties that are invaluable in both the theory and applications of linear algebra, so it should come as no surprise that the

[1] All matrices are assumed to be real for the sake of simplicity, but no logical difficulties are encountered by replacing transpose $(\star)^T$ with conjugate transpose $(\star)^*$ to extend statements in this note to complex matrices.

SVD is key to revealing the exact value $rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T)$. But exactly how this occurs is the point of the following theorem.

Theorem 1. *If* $rank(\mathbf{A}_{\mathbf{m}\times\mathbf{n}}) = r$, *then*

$$rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T) = \begin{cases} r+1 \iff \mathbf{c} \notin R(\mathbf{A}) \textit{ and } \mathbf{d} \notin R\left(\mathbf{A}^T\right) \\ \quad r \quad\;\; \iff \begin{cases} \mathbf{c} \in R(\mathbf{A}) \textit{ xor } \mathbf{d} \in R\left(\mathbf{A}^T\right) \\ \textit{or} \\ \mathbf{c} \in R(\mathbf{A}) \textit{ and } \mathbf{d} \in R\left(\mathbf{A}^T\right) \textit{ and } \alpha \neq 0 \end{cases} \\ r-1 \iff \mathbf{c} \in R(\mathbf{A}) \textit{ and } \mathbf{d} \in R\left(\mathbf{A}^T\right) \textit{ and } \alpha = 0, \end{cases}$$

where $\alpha = 1 + \mathbf{d}^T\mathbf{A}^{\dagger}\mathbf{c}$ *in which* $\mathbf{A}^{\dagger}$ *is the Moore–Penrose pseudo inverse of* $\mathbf{A}$ *and "xor" is the "exclusive or."*

Proof. Let $\mathbf{A} = \mathbf{U}\widehat{\mathbf{D}}\mathbf{V}^T$ be an SVD for $\mathbf{A}$ in which $\widehat{\mathbf{D}} = \begin{pmatrix} \mathbf{D}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, where $\mathbf{D}$ is the diagonal matrix of singular values. If we set

$$\mathbf{x} = \mathbf{U}^T\mathbf{c} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \mathbf{V}^T\mathbf{d} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad \mathbf{x}_1, \mathbf{y}_1 \in \mathbb{R}^{r\times 1}, \tag{6}$$

then

$$rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T) = rank\left(\mathbf{U}\big(\widehat{\mathbf{D}} + \mathbf{x}\mathbf{y}^T\big)\mathbf{V}^T\right) = rank\,(\widehat{\mathbf{D}} + \mathbf{x}\mathbf{y}^T).$$

Since

$$\begin{aligned} rank\,(\widehat{\mathbf{D}} + \mathbf{x}\mathbf{y}^T) + 1 &= rank\begin{pmatrix} \widehat{\mathbf{D}} + \mathbf{x}\mathbf{y}^T & \mathbf{x} \\ \mathbf{0} & -1 \end{pmatrix} = rank\begin{pmatrix} \widehat{\mathbf{D}} & \mathbf{x} \\ \mathbf{y}^T & -1 \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{y}^T & 1 \end{pmatrix} \\ &= rank\begin{pmatrix} \widehat{\mathbf{D}} & \mathbf{x} \\ \mathbf{y}^T & -1 \end{pmatrix}, \end{aligned}$$

it follows that

$$rank\,(\mathbf{A}+\mathbf{c}\mathbf{d}^T) = rank\begin{pmatrix} \widehat{\mathbf{D}} & \mathbf{x} \\ \mathbf{y}^T & -1 \end{pmatrix} - 1 = rank\begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{x}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{x}_2 \\ \mathbf{y}_1^T & \mathbf{y}_2^T & -1 \end{pmatrix} - 1. \tag{7}$$

Furthermore, $\mathbf{d}^T\mathbf{A}^{\dagger}\mathbf{c} = \mathbf{d}^T\mathbf{V}\widehat{\mathbf{D}}^{\dagger}\mathbf{U}^T\mathbf{c} = \mathbf{y}_1^T\mathbf{D}^{-1}\mathbf{x}_1$, so

$$\begin{aligned} &\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\mathbf{y}_1^T\mathbf{D}^{-1} & \mathbf{0} & 1 \end{pmatrix}\begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{x}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{x}_2 \\ \mathbf{y}_1^T & \mathbf{y}_2^T & -1 \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{0} & -\mathbf{D}^{-1}\mathbf{x}_1 \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{x}_2 \\ \mathbf{0} & \mathbf{y}_2^T & -\alpha \end{pmatrix} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix}, \end{aligned}$$

where $\mathbf{Z}$ has the form

$$\mathbf{Z} = \begin{pmatrix} \mathbf{0} & \mathbf{x}_2 \\ \mathbf{y}_2^T & -\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 & \star \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \star \\ \star & \star & \cdots & \star & -\alpha \end{pmatrix}. \tag{8}$$

Hence (7) becomes

$$rank\,(\mathbf{A} + \mathbf{cd}^T) = rank \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix} - 1 = r + rank\,(\mathbf{Z}) - 1. \tag{9}$$

By considering the eight logical possibilities for $\mathbf{x}_2$, $\mathbf{y}_2$, and α being zero or nonzero, it is clear from the structure of $\mathbf{Z}$ in (8) that

$$rank\,(\mathbf{Z}) = \begin{cases} 2 \iff \mathbf{x}_2 \neq \mathbf{0} \text{ and } \mathbf{y}_2 \neq \mathbf{0}, \\ 1 \iff \begin{cases} \mathbf{x}_2 = \mathbf{0} \text{ xor } \mathbf{y}_2 = \mathbf{0}, \\ \text{or} \\ \mathbf{x}_2 = \mathbf{0} \text{ and } \mathbf{y}_2 = \mathbf{0} \text{ and } \alpha \neq 0, \end{cases} \\ 0 \iff \mathbf{x}_2 = \mathbf{0} \text{ and } \mathbf{y}_2 = \mathbf{0} \text{ and } \alpha = 0. \end{cases} \tag{10}$$

If $\mathbf{U}$ and $\mathbf{V}$ are partitioned as $\mathbf{U} = \big[(\mathbf{U}_1)_{m\times r} \,|\, \mathbf{U}_2\big]$ and $\mathbf{V} = \big[(\mathbf{V}_1)_{n\times r} \,|\, \mathbf{V}_2\big]$, then it is the nature of an SVD to have $R(\mathbf{U_1}) = R(\mathbf{A})$ and $R(\mathbf{V_1}) = R(\mathbf{A}^T)$. By using $\mathbf{c} = \mathbf{Ux}$ and $\mathbf{d} = \mathbf{Vy}$ from (6), it follows that

$$\mathbf{x}_2 = \mathbf{0} \iff \mathbf{c} = \mathbf{U}_1\mathbf{x}_1 \in R\,(\mathbf{A}) \quad \text{and} \quad \mathbf{y}_2 = \mathbf{0} \iff \mathbf{d} = \mathbf{V}_1\mathbf{y}_1 \in R\left(\mathbf{A}^T\right),$$

so this together with (9) and (10) yields the desired result. ■

REFERENCES

[1] Campbell, S. L., Meyer, C. D. (2009). *Generalized Inverses of Linear Transformations*. Classics In Applied Mathematics, Number 56. Philadelphia: Society for Industrial and Applied Mathematics, doi.org/10.1137/1.9780898719048

[2] Meyer, C. D. (2000). *Matrix Analysis and Applied Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics.

Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205
meyer@ncsu.edu