Journal of Convex Analysis
Volume 28 (2021), No. 2, 725–728

The Column-Row Factorization of a Matrix

Gilbert Strang
Department of Mathematics, MIT, Cambridge, MA 02139, U.S.A.
gilstrang@gmail.com

Dedicated to Umberto Mosco on the occasion of his 80th birthday.

Received: May 10, 2020
Accepted: May 17, 2020

The active ideas in linear algebra are often expressed by matrix factorizations: $S = Q\Lambda Q^{\mathrm{T}}$ for symmetric matrices (the spectral theorem) and $A = U\Sigma V^{\mathrm{T}}$ for all matrices (singular value decomposition). Far back near the beginning comes $A = LU$ for successful elimination: Lower triangular times upper triangular. This paper is one step earlier, with bases in $A = CR$ for the column space and row space of any matrix – and a proof that column rank = row rank. The echelon form of A and the pseudoinverse A^+ appear naturally. The "proofs" are mostly "observations".

Keywords: Matrix, factorizations, rank, echelon form.

2010 Mathematics Subject Classification: 15A23.

1. Introduction

An introduction is hardly necessary for so short a paper. But I would like to explain the background.

In teaching linear algebra, the course often begins slowly. The idea of a vector space waits until Chapter 3. The highly important topic of singular values is squeezed into the final week or completely omitted. A new plan is needed.

I now start the course in a different way. The multiplication $A\boldsymbol{x}$ produces a combination of the columns of A. All combinations fill the *column space of* A – a key idea to visualize. Simple examples of $A\boldsymbol{x} = \mathbf{0}$ show the idea of linear dependence. Starting with column 1, we create a matrix C with a full set of independent columns – a basis for the column space.

I believe that this "fast start" is also a better start. Every column of A is a combination of the columns of C. Introducing matrix multiplication, *that fact becomes* $A = CR$. We have a natural factorization of A, to be followed by $A = LU$ (elimination) and $A = QR$ (Gram-Schmidt) and $S = Q\Lambda Q^{\mathrm{T}}$ (eigenvalues in the spectral theorem) and $A = U\Sigma V^{\mathrm{T}}$ (singular values in the SVD). The course has a structure that students can follow. A new textbook called "Linear Algebra for Everyone" begins with $A = CR$ [5].

The key point for this paper is that the matrix R in $A = CR$ is already famous. R is the *reduced row echelon form of* A, with any zero rows removed. It has a simple "formula" $R = \begin{bmatrix} I & F \end{bmatrix} P$ which the mechanics of elimination will execute. And

ISSN 0944-6532 / \$ 2.50

it has a "meaning" that is hidden in those row operations on A. R tells us the combinations of independent columns in C, which produce all the columns of A.

This expository paper is dedicated to my friend Umberto Mosco.

2. The Factorization $A = CR$

A is a real matrix with m rows and n columns. Some of those columns might be linear combinations of previous columns. Here is a natural way, working from left to right, to find a complete set of independent columns.

If column 1 is not zero, put it into C.

If column 2 of A is not a multiple of column 1, put it into C.

If column 3 is not a linear combination of columns 1 and 2, put it into C.

Continue.

At the end, C will have r independent columns. Those columns will be a basis for the column space of A. Every column of A is a combination of columns of C, and the coefficients in those combinations go into columns of R:

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} = CR \quad \text{with} \quad r = 2$$

The matrix R contains an r by r identity matrix in the columns that correspond to independent columns of A. Column 3 of R tells us that the columns of A have $\boldsymbol{a}_3 = -\boldsymbol{a}_1 + 2\boldsymbol{a}_2$. A first observation: *$R$ is the "reduced row echelon form" of A, without its $m - r$ zero rows.*

A second observation: *Every row of A is a combination of the rows of R.* That comes directly from the matrix multiplication $A = CR$. In this example row 1 of A equals row 1 of R plus 4 times row 2 of R. The coefficients like 1 and 4 in those linear combinations are in the rows of C. And the rows of R are independent, because an r by r identity matrix is a submatrix of R. Then the rows of R are a basis for the row space of A.

This matrix R is usually computed by row operations on A, to reach the "echelon form". Here R appears after the column basis in C.

A third observation comes from $A = CR = (m \times r)\,(r \times n)$: *The column rank of A equals the row rank of A.* The same number r counts independent columns in C and independent rows in R.

A fourth observation is a "formula" for the reduced row echelon form $R_0 = \mathbf{rref}(A)$. Normally this matrix with $m - r$ zero rows is constructed directly by row operations on A, and C does not appear. A direct description of R_0 could be as follows.

Suppose the basic columns in C are columns $n_1 < n_2 < \cdots < n_r$ of A. The other $n - r$ columns of A are combinations $N = CF$ of those r basic columns (in order). Then the reduced row echelon form of A with $m - r$ zero rows is R_0:

$$R_0 = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} P$$

The permutation P puts the n columns of C and N into their correct order in A, and $I = r \times r$ identity matrix.

Note that $A = C\begin{bmatrix} I & F \end{bmatrix}P = \begin{bmatrix} C & N \end{bmatrix}P$ has the columns of A in their original order, thanks to P. In the formula above, R_0 is constructed directly from A and its uniqueness is clear. Eric Grinberg [1] has invented the name "*gauche basis*" for the columns of C – a brilliant suggestion that reinforces the left-to-right construction of this basis for the column space of A.

Uniqueness of the reduced row echelon form seems to be a moot question when there is an explicit formula for that matrix R_0. The formula cannot be new, but I don't know a reference. Observations 1-3 are definitely not new.

3. A mixing matrix

Here is a variation on the matrix factorization $A = CR$. The matrix C contains actual columns of A but the matrix R does not contain rows of A. For symmetric perfection, we might prefer the matrix R^* containing the r uppermost linearly independent rows taken directly from A. Generally $A = CR^*$ will not be true. To recover a correct factorization of A, we need to include a *mixing matrix* M between C and R^*. Then $A = CMR^*$. (The idea of a mixing matrix has become widespread in numerical linear algebra. The symbol U is often chosen instead of M, but U is needed in the important factorizations LU and $U\Sigma V^{\mathrm{T}}$. We would like to nominate the letter M and the adjective *mixing*.)

Does M have a simple formula? *Yes.* M^{-1} is the r by r submatrix at the intersection of the r columns of C with the r rows of R^*. If those happen to be the first r columns and the first r rows, it is easy to see that MR^* will produce the familiar r by r identity matrix that begins the reduced row echelon form R. Then $A = CMR^*$ is identified with $A = CR$.

4. The pseudoinverse

Factorizations like $A = CR$ are familiar to algebraists. It is not surprising that they connect to other constructions. An example in linear algebra is the *pseudoinverse of* A.

We write the pseudoinverse as A^+. It inverts the mapping (multiplication by A) from row space to column space. The pseudoinverse is *zero* on the nullspace of A^{T}. Thus it inverts A where inversion is possible: $AA^+A = A$.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A^+ = (U\Sigma V^{\mathrm{T}})^+ = V\Sigma^+U^{\mathrm{T}}$$

The pseudoinverse connects perfectly to the rank r factors in $A = CR$. The pseudoinverse of C is its left inverse $C^+ = (C^{\mathrm{T}}C)^{-1}C^{\mathrm{T}}$. The pseudoinverse of R is its right inverse $R^+ = R^{\mathrm{T}}(RR^{\mathrm{T}})^{-1}$. Then the pseudoinverse of $A = CR$ is $A^+ = R^+C^+$, because all ranks are equal to r.

The early papers of T. N. E Greville in the first issues of SIAM Review [1, 2] foreshadowed the ideas in many later papers, including this one. The work of E. M. Moore is referenced, as well as the papers of Penrose which led to the original name "Moore-Penrose inverse" for A^+. And a third author Bjerkammar seems to deserve credit too ...

References

[1] T. N. E. Greville: *The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations*, SIAM Review 1 (1959) 38–43.

[2] T. N. E. Greville: *Some applications of the pseudoinverse of a matrix*, SIAM Review 2 (1960) 15–22.

[3] E. Grinberg: *A Gauche perspective on the row reduced echelon form and its uniqueness*, arXiv: 2005.06275 (2020).

[4] G. Strang: *Multiplying and factoring matrices*, Amer. Math. Monthly 125 (2018) 223–230.

[5] G. Strang: *Linear Algebra for Everyone*, Wellesley-Cambridge Press, Wellesley (2020).

[6] G. Strang, D. Drucker, A. Lin: *Rank-revealing factorizations*, in preparation (2020).

[7] G. Strang, S. Lee, *Row reduction of a matrix and $A = CaB$*, Amer. Math. Monthly 107 (2000) 681–688.