Elimination and Factorization $A = CR$

Gilbert Strang, MIT

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 11 & 17 \\ 3 & 7 & 37 & 57 \\ 4 & 9 & 48 & 74 \end{bmatrix} \to \boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 3 & 5 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

1. Rows $1, 2$ of $\boldsymbol{Z}$ (call them $\boldsymbol{R}$) are a basis for the row space of A.

2. Columns $1, 2$ of $\boldsymbol{A}$ (call them $\boldsymbol{C}$) are a basis for the column space of A.

3. The nullspace of $\boldsymbol{Z}$ equals the nullspace of $\boldsymbol{A}$ (orthogonal to the same row space).

Elimination factors $\boldsymbol{A}$ into $\boldsymbol{C}$ times $\boldsymbol{R}$

$=(\boldsymbol{m}\times\boldsymbol{r})$ times $(\boldsymbol{r}\times\boldsymbol{n})$

$$\boldsymbol{A}=\begin{bmatrix}1 & 2 & 11 & 17\\ 3 & 7 & 37 & 57\\ 4 & 9 & 48 & 74\end{bmatrix}=\begin{bmatrix}1 & 2\\ 3 & 7\\ 4 & 9\end{bmatrix}\begin{bmatrix}1 & 0 & 3 & 5\\ 0 & 1 & 4 & 6\end{bmatrix}=\boldsymbol{CR}$$

$\boldsymbol{C}$ has full column rank $r=2$, $\boldsymbol{R}$ has full row rank $r=2$.
$\boldsymbol{A}=\boldsymbol{CR}$ leads to the first great theorem in linear algebra
Column rank equals row rank for every matrix $\boldsymbol{A}$

Suppose A has r independent columns: rank $= r$
Put the first r independent columns of A into C

Then the other $n - r$ columns of A must be combinations CF of those independent columns in C.

The row factor is $R = \left[I \quad F \quad P \right.$, with r independent rows.

$$A = CR = \left[C \quad CF \quad P \right.$$
$$= \left[\text{Indep cols} \quad \text{Dep cols} \quad \text{Permute cols} \right.$$

If the r independent columns come first in A,
that permutation matrix will be $\boldsymbol{P} = \boldsymbol{I}$.
Otherwise we need P to permute the columns of C
and CF into correct position in A.

$\boldsymbol{P}$ exchanges columns 2 and 3 :

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} P$$

$$= \boldsymbol{CR}.$$

The essential information in $\boldsymbol{Z} = \textbf{rref}(\boldsymbol{A})$ is the list of r independent columns of $\boldsymbol{A}$, and the matrix $\boldsymbol{F}$ (r by $n-r$) that combines those independent columns to give the $n-r$ dependent columns $\boldsymbol{CF}$ in $\boldsymbol{A}$. This uniquely defines $\textbf{rref}(\boldsymbol{A})$.

These **row operations** put $\boldsymbol{A}$ into its reduced row echelon form $\boldsymbol{Z}$

(a) Subtract a multiple of one row from another row (below or above)
(b) Exchange two rows
(c) Divide a row by its first nonzero entry

Elimination reduces $\boldsymbol{A}$ to $\boldsymbol{Z} = \textbf{rref}(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \boldsymbol{P}$

Column by column (left to right) construct $\boldsymbol{Z} = \textbf{rref}(\boldsymbol{A})$.
After elimination on k columns, that first part of
the matrix is in its own **rref** form.
The next column has an upper part $\boldsymbol{u}$ and a lower part $\boldsymbol{\ell}$:

First $\boldsymbol{k+1}$ columns $\begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{F}_k \\ 0 & 0 \end{bmatrix} \boldsymbol{P}_k$ followed by $\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{\ell} \end{bmatrix}$

The big question is: **Does this new column $\boldsymbol{k+1}$ join with $\boldsymbol{I}_k$ or $\boldsymbol{F}_k$?**

If $\boldsymbol{\ell}$ is all zeros, the new column is **dependent** on the first k columns. Then **$\boldsymbol{u}$ joins with $\boldsymbol{F_k}$** to produce $\boldsymbol{F_{k+1}}$ in the next step to column $k+2$.

If $\boldsymbol{\ell}$ is not all zero, the new column is **independent** of the first k columns.
Pick any nonzero in $\boldsymbol{\ell}$ as the *pivot*.
Move that pivot row of A to the top of $\boldsymbol{\ell}$.
Use that row to zero out all the rest of column $k+1$.
The new column $k+1$ joins with $\boldsymbol{I_k}$ to produce $\boldsymbol{I_{k+1}}$.

Elimination tells us the **first r independent columns of A**. Those are the columns of C. The row space is not changed! Then its orthogonal complement (**the nullspace of A**) is not changed.
Each column of CF tells us how a dependent column of A is a combination of the independent columns in C.

Key point: The $n - r$ columns of F are telling us $n - r$ solutions to $Ax = 0$.

The **nullspace of** $\boldsymbol{A}$ is easiest to see by example.

$$\begin{array}{rcl} x_1+2x_2+11x_3+17x_4=0 & \text{reduces} & \boldsymbol{x_1+3x_3+5x_4=0} \\ 3x_1+7x_2+37x_3+57x_4=0 & \text{to} & \boldsymbol{x_2+4x_3+6x_4=0} \end{array}$$

Solution with $x_3=1$ & $x_4=0$ is $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{-3} & \boldsymbol{-4} & 1 & 0\end{bmatrix}^{\mathrm{T}}$.

Solution with $x_3=0$ & $x_4=1$ is $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{-5} & \boldsymbol{-6} & 0 & 1\end{bmatrix}^{\mathrm{T}}$.

Those solutions are the two columns of $\boldsymbol{X}$ in $\boldsymbol{AX}=\boldsymbol{0}$.

Find the **nullspace of $\boldsymbol{A}$**

$\boldsymbol{A} = \boldsymbol{C}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \end{bmatrix}\boldsymbol{P}$ multiplies the nullspace matrix

$\boldsymbol{X} = \boldsymbol{P}^{\mathrm{T}}\begin{bmatrix} -\boldsymbol{F} \\ \boldsymbol{I}_{n-r} \end{bmatrix}$ to produce

$\boldsymbol{AX} = -\boldsymbol{CF} + \boldsymbol{CF} = \mathbf{0}.$

Each column of $\boldsymbol{X}$ solves $\boldsymbol{Ax} = \mathbf{0}$ (note that $\boldsymbol{PP}^{\mathrm{T}} = \boldsymbol{I}$). Each dependent column of $\boldsymbol{A}$ is a combination of the independent columns in $\boldsymbol{C}$.

Gauss-Jordan elimination leading to $A = CR$ is less efficient than the Gauss process that directly solves $A\boldsymbol{x} = \boldsymbol{b}$. Gauss stops at a **triangular system $U\boldsymbol{x} = \boldsymbol{c}$**: back substitution produces $\boldsymbol{x}$. Gauss-Jordan has the extra cost of eliminating upwards. If we only want to solve equations, Gauss is faster than $A = CR$.

Block elimination Suppose the matrix $\boldsymbol{W}$ in the first r rows and columns of A is invertible. Then *elimination takes all its instructions from* W !

$\boldsymbol{W}$ will change to $\boldsymbol{I}$. This identifies $\boldsymbol{F}$ as $\boldsymbol{W}^{-1}\boldsymbol{H}$.

And the last $m - r$ rows will become zero rows.

$$\textbf{Block elimination} \quad \boldsymbol{A} = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{H} \\ \boldsymbol{J} & \boldsymbol{K} \end{bmatrix} \text{ reduces to } \begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \textbf{rref}(\boldsymbol{A})$$

In general that first r by r block might not be invertible. But elimination will find W. We can move W to the upper left corner by row and column permutations $\boldsymbol{P_r}$ and $\boldsymbol{P_c}$. Then the full expression of block elimination is

$$\boldsymbol{P}_r \boldsymbol{A} \boldsymbol{P}_c = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{H} \\ \boldsymbol{J} & \boldsymbol{K} \end{bmatrix} \to \begin{bmatrix} \boldsymbol{I} & \boldsymbol{W}^{-1}\boldsymbol{H} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

Interesting point. Since A has rank r, we know that A has r independent rows and r independent columns. Suppose those rows are in a submatrix B and those columns are in a submatrix C. Is it always true that the r by r "intersection" W of those rows B with those columns C will be **invertible** ?

Yes ! The intersection of $\boldsymbol{r}$ independent rows of $\boldsymbol{A}$ with $\boldsymbol{r}$ independent columns **of $\boldsymbol{A}$** is invertible.

To appear in Mathematics Magazine 12/2024 (MAA Journal)
This paper has appeared in arXiv.org/abs/2304.02659

Introduction to Linear Algebra, 6th edition

Wellesley-Cambridge Press **gilstrang@gmail.com**

Resource: A Vision of Linear Algebra
Gilbert Strang