Linear Substitutions and Matrix Multiplication

This note interprets matrix multiplication and related concepts in terms of the composition of linear substitutions.

Composition and multiplication We start from the linear substitution (cf. Example 3 on page 5 of Anton–Rorres)

\[
\begin{align*}
\begin{cases}
  u = x + y + 2z \\
v = 2x + 4y - 3z \\
w = 3x + 6y - 5z
\end{cases}
\text{ with matrix } A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix}
\]
\]

(Linear substitutions are usually not allowed to have constant terms.) Then we substitute in turn linear expressions for each of \( x, y, \) and \( z \):

\[
\begin{align*}
\begin{cases}
  x = s + 2t \\
y = 3s - t \\
z = -s + t
\end{cases}
\text{ with matrix } B = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ -1 & 1 \end{bmatrix}
\end{align*}
\]

We see that \( u, v, \) and \( w \) then become linear expressions in \( s \) and \( t \),

\[
\begin{align*}
\begin{cases}
  u = x + y + 2z &= s + 2t + 3s - t - 2s + 2t = 2s + 3t \\
v = 2x + 4y - 3z &= 2s + 4t + 12s - 4t + 3s - 3t = 17s - 3t \\
w = 3x + 6y - 5z &= 3s + 6t + 18s - 6t + 5s - 5t = 26s - 5t
\end{cases}
\end{align*}
\]

This is the composite linear substitution of the linear substitutions represented by equations (1) and (2). We observe that its matrix \( D \) is just the product matrix \( AB \),

\[
D = AB = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 17 & -3 \\ 26 & -5 \end{bmatrix}
\]

In other words, in terms of the vectors

\[
s = \begin{bmatrix} s \\ t \end{bmatrix}, \quad x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad u = \begin{bmatrix} u \\ v \\ w \end{bmatrix}
\]

we have

\[
u = Ax = A(Bs) = (AB)s = Ds
\]

Associativity Suppose we introduce another linear substitution \( s = Cp \), where we find that there will be no need to make \( p \) explicit. Then we can express \( x \) in terms of \( p \) by the linear substitution

\[
x = Bs = B(Cp) = Ep,
\]

where \( E = BC \) denotes another product matrix.

We now have two ways to express \( u \) in terms of \( p \). We can use equation (3) to go by way of \( s \),

\[
u = Ds = D(Cp) = (DC)p
\]
Or we can use equations (1) and (4) to go by way of $x$,

$$u = Ax = A(Ep) = (AE)p$$

We therefore conclude, without doing any real work, that $AE = DC$, that is, $A(BC) = (AB)C$, which is the *associative law*. Alternatively, we could express this diagrammatically as

$$p \xrightarrow{C} s \xrightarrow{B} x \xrightarrow{A} u$$

which indicates that $(AB)C$ and $A(BC)$ are really the same. We may simply write $u = ABCp$, with no parentheses.

**The identity**  The *identity* linear substitution

$$\begin{align*}
x &= x \\
y &= y \\
z &= z
\end{align*}$$

whose matrix is

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

clearly preserves everything. Its matrix $I$ (in this case $I_3$) is therefore called the *identity* matrix. Generally, there is an $n \times n$ identity matrix $I_n$ for each $n$. (But it has to be square.) From this point of view, it is obvious that $I_mF = F = FI_n$ for any $m \times n$ matrix $F$.

**The zero**  In contrast, the *zero* linear substitution

$$\begin{align*}
y_1 &= 0 \\
y_2 &= 0 \\
y_3 &= 0 \\
y_4 &= 0
\end{align*}$$

destroys everything in sight by setting all variables zero. Its matrix is the *zero matrix*, with all entries 0. The zero matrix comes in all sizes $m \times n$. Applying other linear substitutions before or after the zero linear substitution will still leave all variables set to 0. So in matrix language, $G0 = 0$ and $0G = 0$ for any matrix $G$, whenever the sizes of these zero matrices allow the products to be defined.