References

[1] Leonard E. Dickson, History of the Theory of Numbers, vol. 1, Chelsea, New York, 1952.
[2] Christopher Hooley, On Artin's conjecture, J. Reine Angew. Math., 225 (1967) 209–220.

On the Inverse of the Sum of Matrices

KENNETH S. MILLER
Riverside Research Institute
80 West End Avenue
New York, NY 10023

If $\mathbf{G}$ and $\mathbf{H}$ are arbitrary nonsingular square matrices of the same dimension, then the inverse of their product $\mathbf{GH}$ is well known to be $\mathbf{H}^{-1}\mathbf{G}^{-1}$. If $\mathbf{G}$ and $\mathbf{G}+\mathbf{H}$ are nonsingular, then the problem of finding a simple expression for the inverse of $\mathbf{G}+\mathbf{H}$, say in terms of $\mathbf{G}^{-1}$, is more difficult. This is the problem we wish to address. In the theorem below we shall establish a recursive form for $(\mathbf{G}+\mathbf{H})^{-1}$.

The key result in proving this theorem is a fundamental Lemma which states that if $\mathbf{H}$ has rank one, then

$$(\mathbf{G}+\mathbf{H})^{-1}=\mathbf{G}^{-1}-\frac{1}{1+g}\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1} \tag{1}$$

where $g=\operatorname{tr}\mathbf{H}\mathbf{G}^{-1}$. For example, if

$$\mathbf{A}=\begin{bmatrix} a & 1 & 1\\ 1 & a & 1\\ 1 & 1 & a\end{bmatrix}, \qquad a\neq 1,-2$$

is a 3×3 nonsingular matrix, then we may write $\mathbf{A}=\mathbf{G}+\mathbf{H}$ where $\mathbf{G}$ is $(a-1)$ times the identity matrix $\mathbf{I}$ and $\mathbf{H}$ is a 3×3 matrix with ones everywhere (and hence of rank one). Since $\operatorname{tr}\mathbf{H}\mathbf{G}^{-1}=3/(a-1)$, equation (1) yields

$$\mathbf{A}^{-1}=\frac{1}{a-1}\left[\mathbf{I}-\frac{1}{a+2}\mathbf{H}\right]$$
$$=\frac{1}{(a-1)(a+2)}\begin{bmatrix} a+1 & -1 & -1\\ -1 & a+1 & -1\\ -1 & -1 & a+1\end{bmatrix}.$$

That the above formula indeed represents the inverse of $\mathbf{A}$ may be verified by a direct calculation.

Our program will be as follows. First we shall prove (1). Then we shall show that any matrix may be decomposed into the sum of matrices of rank one. This result will enable us to iteratively apply the Lemma to obtain our main theorem. We also shall consider some ancillary results, including some simple generalizations to the inverse of sums of Kronecker products of matrices.

It is appropriate to begin our analysis by considering matrices of rank one. Suppose, then, that $\mathbf{E}$ is a square matrix of rank one. Then all except possibly one eigenvalue of $\mathbf{E}$ is zero. Since the sum of the eigenvalues of $\mathbf{E}$ is the trace of $\mathbf{E}$, we see that the remaining eigenvalue is $\operatorname{tr}\mathbf{E}$. Matrices of rank one may be constructed by beginning with two nonzero column vectors $\mathbf{u}$ and $\mathbf{v}$ of the same dimension. Since $\mathbf{u}$ and $\mathbf{v}$ have rank one, the matrix $\mathbf{E}=\mathbf{u}\mathbf{v}'$ has rank one. (We shall use primes to indicate transposes of vectors and matrices. In particular, $\mathbf{v}'$ is a row vector.) Conversely, if $\mathbf{E}$ has rank one, then there exist nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ such that $\mathbf{E}=\mathbf{u}\mathbf{v}'$ [5, page 92]. We give some examples to indicate the usefulness of this representation.

EXAMPLE 1. (Norm of a matrix) If $\mathbf{A}$ is any square matrix, then the **norm** $|\mathbf{A}|$ of $\mathbf{A}$ is the square root of the largest eigenvalue of $\mathbf{A}'\mathbf{A}$. Now suppose $\mathbf{E}$ is a square matrix of rank one, and we write $\mathbf{E}=\mathbf{uv}'$ where $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors. Then $\mathbf{E}'\mathbf{E}=\mathbf{vu}'\mathbf{uv}'=|\mathbf{u}|^2\mathbf{vv}'$ and the only nonzero eigenvalue of $\mathbf{E}'\mathbf{E}$ is $|\mathbf{u}|^2\operatorname{tr}\mathbf{vv}'=|\mathbf{u}|^2|\mathbf{v}|^2$. Hence

$$|\mathbf{E}|=|\mathbf{u}||\mathbf{v}|. \tag{2}$$

As a corollary we deduce from the Schwarz inequality $|\mathbf{u}'\mathbf{v}|\leqslant|\mathbf{u}||\mathbf{v}|$ that

$$|\operatorname{tr}\mathbf{E}|\leqslant|\mathbf{E}|. \tag{3}$$

This result will be used later.

EXAMPLE 2. (Reproducing property) Let $\mathbf{A}$ be a square matrix and $\mathbf{E}$ a square matrix of rank one. Then if we write $\mathbf{E}=\mathbf{uv}'$ where $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, we have the interesting and useful reproducing property

$$\mathbf{EAE}=\mathbf{uv}'\mathbf{Auv}'=\beta\mathbf{uv}'=\beta\mathbf{E} \tag{4}$$

where β is the bilinear form $\mathbf{v}'\mathbf{Au}$. We also may write

$$\beta=\operatorname{tr}\mathbf{v}'\mathbf{Au}=\operatorname{tr}\mathbf{uv}'\mathbf{A}=\operatorname{tr}\mathbf{EA}.$$

Thus (4) may be written as

$$\mathbf{EAE}=(\operatorname{tr}\mathbf{EA})\mathbf{E} \tag{5}$$

and no explicit mention of the vectors $\mathbf{u}$ and $\mathbf{v}$ appears. As a corollary we see that

$$(\mathbf{EA})^n=\beta^{n-1}\mathbf{EA}\quad\text{and}\quad(\mathbf{AE})^n=\beta^{n-1}\mathbf{AE} \tag{6}$$

for all positive integers n.

Equation (4) is the key formula in proving the fundamental Lemma mentioned at the beginning of this paper. Suppose, then, that $\mathbf{G}$ and $\mathbf{G}+\mathbf{E}$ are nonsingular matrices where $\mathbf{E}$ has rank one. We may write $\mathbf{G}+\mathbf{E}=(\mathbf{I}+\mathbf{EG}^{-1})\mathbf{G}$, and since $\mathbf{G}$ is nonsingular, the matrix $\mathbf{EG}^{-1}$ has rank one. Hence $1+\operatorname{tr}\mathbf{EG}^{-1}$ is an eigenvalue of $\mathbf{I}+\mathbf{EG}^{-1}$, the remaining eigenvalues all being one. Thus we see that $\mathbf{G}+\mathbf{E}$ is nonsingular if and only if $\operatorname{tr}\mathbf{EG}^{-1}\neq-1$.

We search for an inverse of $\mathbf{G}+\mathbf{E}$ in the form $\mathbf{G}^{-1}-\nu\mathbf{G}^{-1}\mathbf{EG}^{-1}$ where ν is a scalar. (A possible motivation to justify this choice stems from the following argument. Let $f(\zeta)$ be continuous on $[0,x]$ and differentiable on $(0,x)$. Let $x<a$. Then by the Law of the Mean, $f(x)=f(0)+f'(\lambda a)x$ where $0<\lambda<1$. Now let $f(\zeta)=(\zeta+a)^{-1}$. Then we may write

$$(x+a)^{-1}=a^{-1}-\nu a^{-1}xa^{-1}$$

where $\nu=(1+\lambda)^{-2}$. Identify x with $\mathbf{E}$ and a with $\mathbf{G}$.)

If $\mathbf{G}^{-1}-\nu\mathbf{G}^{-1}\mathbf{EG}^{-1}$ is to be the inverse of $\mathbf{G}+\mathbf{E}$, their product

$$(\mathbf{G}+\mathbf{E})(\mathbf{G}^{-1}-\nu\mathbf{G}^{-1}\mathbf{EG}^{-1})=\mathbf{I}-\nu\mathbf{EG}^{-1}+\mathbf{EG}^{-1}-\nu\mathbf{EG}^{-1}\mathbf{EG}^{-1}$$

must be the identity matrix. This in turn implies that

$$\nu\mathbf{EG}^{-1}-\mathbf{EG}^{-1}+\nu(\mathbf{EG}^{-1})^2=\mathbf{0}. \tag{7}$$

By (6), $(\mathbf{EG}^{-1})^2=g\mathbf{EG}^{-1}$ where $g=\operatorname{tr}\mathbf{EG}^{-1}$ and (7) may be written in the form

$$(\nu-1+\nu g)\mathbf{EG}^{-1}=\mathbf{0}. \tag{8}$$

A sufficient condition for (8) to be valid is to have $\nu-1+\nu g=0$ or

$$\nu=\frac{1}{1+g}.$$

(Since $\mathbf{G}+\mathbf{E}$ is nonsingular, we have seen that $1+g\neq 0$.) Thus we have proved:

LEMMA. *Let* $\mathbf{G}$ *and* $\mathbf{G}+\mathbf{E}$ *be nonsingular matrices where* $\mathbf{E}$ *is a matrix of rank one. Let* $g=\operatorname{tr}\mathbf{EG}^{-1}$. *Then* $g\neq-1$ *and*

$$\mathbf{G}^{-1}-\frac{1}{1+g}\mathbf{G}^{-1}\mathbf{EG}^{-1}$$

is the inverse of $\mathbf{G}+\mathbf{E}$.

The above equation is essentially the Sherman-Morrison formula (see [1, page 161]).

Before continuing to the general case of finding the inverse of $\mathbf{G}+\mathbf{H}$ where $\mathbf{H}$ is not necessarily of rank one, let us show the relation of this Lemma to the Neumann series expansion of a matrix.

EXAMPLE 3. (Neumann series) If $\mathbf{P}$ is a square matrix and $|\mathbf{P}|<1$, then $(\mathbf{I}-\mathbf{P})^{-1}$ has the Neumann series expansion

$$(\mathbf{I}-\mathbf{P})^{-1}=\mathbf{I}+\mathbf{P}+\mathbf{P}^2+\cdots+\mathbf{P}^n+\cdots \tag{9}$$

(see [5, page 186]). Now let us suppose that $\mathbf{P}$ also has rank one. Then by (6), $\mathbf{P}^n=\alpha^{n-1}\mathbf{P}$, $n=1,2,\ldots$ where $\alpha=\operatorname{tr}\mathbf{P}$. Thus we may write (9) as

$$(\mathbf{I}-\mathbf{P})^{-1}=\mathbf{I}+(1+\alpha+\cdots+\alpha^{n-1}+\cdots)\mathbf{P}. \tag{10}$$

The series on the right-hand side of (10) is a geometric series with ratio α, and by (3) $|\alpha|=|\operatorname{tr}\mathbf{P}|\leqslant|\mathbf{P}|<1$. Hence we may sum the series to obtain

$$(\mathbf{I}-\mathbf{P})^{-1}=\mathbf{I}+\frac{1}{1-\alpha}\mathbf{P}.$$

This formula is just a special case of the Lemma. One may consider the above identity for $(\mathbf{I}-\mathbf{P})^{-1}$ to be an "analytic continuation" of (9).

Now let us return to our main problem. We wish to find the inverse of $\mathbf{G}+\mathbf{H}$ where $\mathbf{G}$ and $\mathbf{G}+\mathbf{H}$ are nonsingular. Our basic argument in the Theorem below is to decompose $\mathbf{H}$ into the sum of matrices of rank one and iteratively apply the Lemma. It is known that if $\mathbf{H}$ has positive rank r, then we may write $\mathbf{H}$ in the form

$$\mathbf{H}=\mathbf{E}_1+\mathbf{E}_2+\cdots+\mathbf{E}_r \tag{11}$$

where each $\mathbf{E}_k$, $1\leqslant k\leqslant r$, has rank one [5, page 93]. (This decomposition is not unique.) Thus we may write

$$\mathbf{G}+\mathbf{H}=\mathbf{G}+\mathbf{E}_1+\cdots+\mathbf{E}_r.$$

If we are to recursively apply the Lemma, we need $\mathbf{G}+\mathbf{E}_1+\cdots+\mathbf{E}_k$ to be nonsingular for all k. This will not necessarily be true for an arbitrary decomposition (11) of $\mathbf{H}$. However, we can show that there *does* exist a decomposition (11) such that each of the "partial sums" $\mathbf{C}_{k+1}=\mathbf{G}+\mathbf{E}_1+\cdots+\mathbf{E}_k$ is nonsingular for $k=1,\ldots,r$.

To prove this contention we first note that since $\mathbf{G}+\mathbf{H}=(\mathbf{I}+\mathbf{H}\mathbf{G}^{-1})\mathbf{G}$ is nonsingular, no eigenvalue of $\mathbf{H}\mathbf{G}^{-1}$ can be -1 (as we observed in the proof of the Lemma.) Now let $\mathbf{Q}$ be a nonsingular matrix such that $\mathbf{J}=\mathbf{Q}(\mathbf{H}\mathbf{G}^{-1})\mathbf{Q}^{-1}$ is the Jordan normal form of $\mathbf{H}\mathbf{G}^{-1}$, and let

$$\mathbf{J}=\mathbf{F}_1+\mathbf{F}_2+\cdots+\mathbf{F}_r$$

where the kth row of $\mathbf{F}_k$ is the kth row of $\mathbf{J}$, and the remaining rows of $\mathbf{F}_k$ are all zero. Then every $\mathbf{F}_j$ has rank one and $\mathbf{I}+\mathbf{F}_1+\cdots+\mathbf{F}_k$ is nonsingular for $k=1,\ldots,r$. Thus

$$\left[\mathbf{Q}^{-1}(\mathbf{I}+\mathbf{F}_1+\cdots+\mathbf{F}_k)\mathbf{Q}\right]\mathbf{G}=\mathbf{G}+\mathbf{Q}^{-1}\mathbf{F}_1\mathbf{Q}\mathbf{G}+\cdots+\mathbf{Q}^{-1}\mathbf{F}_k\mathbf{Q}\mathbf{G}$$

is also nonsingular for $k=1,\ldots,r$. Since each $\mathbf{Q}^{-1}\mathbf{F}_j\mathbf{Q}\mathbf{G}$ has rank one, let $\mathbf{E}_j=\mathbf{Q}^{-1}\mathbf{F}_j\mathbf{Q}\mathbf{G}$. Then $\mathbf{C}_{k+1}=\mathbf{G}+\mathbf{E}_1+\cdots+\mathbf{E}_k$ is nonsingular for $k=1,\ldots,r$ and $\mathbf{C}_{r+1}=\mathbf{G}+\mathbf{H}$ where $\mathbf{H}=\mathbf{E}_1+\mathbf{E}_2+\cdots+\mathbf{E}_r$.

We are now in a position to give a precise statement of our main result.

THEOREM. *Let* $\mathbf{G}$ *and* $\mathbf{G}+\mathbf{H}$ *be nonsingular matrices and let* $\mathbf{H}$ *have positive rank* r. *Let* $\mathbf{H}=\mathbf{E}_1+\mathbf{E}_2+\cdots+\mathbf{E}_r$ *where each* $\mathbf{E}_k$ *has rank one and* $\mathbf{C}_{k+1}=\mathbf{G}+\mathbf{E}_1+\cdots+\mathbf{E}_k$ *is nonsingular for* $k=1,\ldots,r$. *Then if* $\mathbf{C}_1=\mathbf{G}$,

$$\mathbf{C}_{k+1}^{-1}=\mathbf{C}_k^{-1}-\nu_k\mathbf{C}_k^{-1}\mathbf{E}_k\mathbf{C}_k^{-1},\quad k=1,\ldots,r$$

where

$$\nu_k = \frac{1}{1 + \operatorname{tr} \mathbf{C}_k^{-1}\mathbf{E}_k}.$$

In particular

$$(\mathbf{G}+\mathbf{H})^{-1} = \mathbf{C}_r^{-1} - \nu_r \mathbf{C}_r^{-1}\mathbf{E}_r\mathbf{C}_r^{-1}.$$

To prove this result we first write $\mathbf{C}_2 = \mathbf{C}_1 + \mathbf{E}_1 = \mathbf{G} + \mathbf{E}_1$ and recall that $\mathbf{G}$ and $\mathbf{C}_2$ are nonsingular. Then by the Lemma,

$$\mathbf{C}_2^{-1} = \mathbf{G}^{-1} - \nu_1 \mathbf{G}^{-1}\mathbf{E}_1\mathbf{G}^{-1} \tag{12}$$

and we have calculated $\mathbf{C}_2^{-1}$ in terms of $\mathbf{G}^{-1}$. Now $\mathbf{C}_3 = \mathbf{G} + \mathbf{E}_1 + \mathbf{E}_2 = \mathbf{C}_2 + \mathbf{E}_2$. Hence since $\mathbf{C}_2$ and $\mathbf{C}_3$ are nonsingular, we again may invoke the Lemma to write $\mathbf{C}_3^{-1}$ in terms of $\mathbf{C}_2^{-1}$, viz.:

$$\mathbf{C}_3^{-1} = \mathbf{C}_2^{-1} - \nu_2 \mathbf{C}_2^{-1}\mathbf{E}_2\mathbf{C}_2^{-1}.$$

But $\mathbf{C}_2^{-1}$ is known from (12). If we continue this process r times (where r is the rank of $\mathbf{H}$) we obtain

$$\mathbf{C}_{r+1}^{-1} = \mathbf{C}_r^{-1} - \nu_r \mathbf{C}_r^{-1}\mathbf{E}_r\mathbf{C}_r^{-1}.$$

But $\mathbf{C}_{r+1} = \mathbf{G} + \mathbf{H}$, and thus our Theorem is proved.

As an illustration let us consider the problem of finding the inverse of $\mathbf{I}+\mathbf{H}$ where $\mathbf{I}$ is the identity matrix and $\mathbf{H}$ has rank two. Applying the above theorem with two (the rank of $\mathbf{H}$) iterations, some algebra will show that

$$(\mathbf{I}+\mathbf{H})^{-1} = \mathbf{I} - \frac{1}{a+b}(a\mathbf{H} - \mathbf{H}^2) \tag{13}$$

where $a = 1 + \operatorname{tr}\mathbf{H}$ and $2b = (\operatorname{tr}\mathbf{H})^2 - \operatorname{tr}\mathbf{H}^2$. Equation (13) also is valid if $\mathbf{H}$ has rank one. For in this case $\mathbf{H}^2 = (\operatorname{tr}\mathbf{H})\mathbf{H}$ by (5) and hence $\operatorname{tr}\mathbf{H}^2 = (\operatorname{tr}\mathbf{H})^2$. Thus $b = 0$ and (13) becomes

$$(\mathbf{I}+\mathbf{H})^{-1} = \mathbf{I} - \frac{1}{a}[a\mathbf{H} - (\operatorname{tr}\mathbf{H})\mathbf{H}] = \mathbf{I} - \frac{1}{a}\mathbf{H}$$

since $\operatorname{tr}\mathbf{H} = a - 1$. But this is just the Lemma with $\mathbf{G} = \mathbf{I}$ and $\mathbf{E} = \mathbf{H}$.

We also may exploit the above theorem to find the determinant of $\mathbf{G}+\mathbf{H}$ in terms of the ν_k.

EXAMPLE 4. (Determinant of $\mathbf{G}+\mathbf{H}$) From the Theorem we see that $\mathbf{C}_{k+1} = \mathbf{C}_k + \mathbf{E}_k = \mathbf{C}_k(\mathbf{I} + \mathbf{C}_k^{-1}\mathbf{E}_k)$ and hence $\det \mathbf{C}_{k+1} = (\det \mathbf{C}_k)\det(\mathbf{I} + \mathbf{C}_k^{-1}\mathbf{E}_k)$. But $\mathbf{C}_k^{-1}\mathbf{E}_k$ has rank one. Thus $\det(\mathbf{I} + \mathbf{C}_k^{-1}\mathbf{E}_k) = 1 + \operatorname{tr}\mathbf{C}_k^{-1}\mathbf{E}_k = \nu_k^{-1}$. Hence $\det \mathbf{C}_{k+1} = \nu_k^{-1}(\det \mathbf{C}_k)$, $k = 1, \ldots, r$ and inductively

$$\det \mathbf{C}_{r+1} = \frac{1}{\nu_1\nu_2\cdots\nu_r}\det \mathbf{C}_1.$$

But $\mathbf{C}_{r+1} = \mathbf{G}+\mathbf{H}$ and $\mathbf{C}_1 = \mathbf{G}$. Hence $(\nu_1\nu_2\cdots\nu_r)\det(\mathbf{G}+\mathbf{H}) = \det \mathbf{G}$.

In many practical problems, for example, in the theory of Kalman filtering (see [4], [6], [8]), one wishes to find the inverse of $\mathbf{G}+\mathbf{H}$ for various matrices $\mathbf{H}$ when $\mathbf{G}^{-1}$ is known. The usefulness of the above formulas in such cases is readily apparent. Frequently in such problems $\mathbf{G}$ is positive definite and $\mathbf{H}$ is diagonal and nonnegative definite. In such cases $\mathbf{G}+\mathbf{H}$ is always nonsingular and the decomposition of $\mathbf{H}$ into matrices of rank one is trivial. It is also to be noted that the recursive form is especially convenient for computer utilization. The techniques also may be applied to the problem of finding the inverse of $\mathbf{G}$ itself. For example, for any square matrix $\mathbf{G}$ we may choose a matrix $\mathbf{G}_1$ and write $\mathbf{G} = \mathbf{G}_1 + (\mathbf{G} - \mathbf{G}_1)$. If $\mathbf{G}_1$ is chosen to be nonsingular, then we may determine $\mathbf{G}^{-1}$ in terms of $\mathbf{G}_1^{-1}$. In particular $\mathbf{G}_1$ may be chosen to be the identity matrix.

Before continuing we consider a specific application of the above Theorem. Let

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & \lambda & \lambda^2 \\ \lambda & 1 & \lambda \\ \lambda^2 & \lambda & 1 \end{bmatrix}, \quad 0 < \lambda < 1$$

be a 3×3 Markoffian matrix. We shall find $\mathbf{\Lambda}^{-1}$. We first write $\mathbf{\Lambda}=\mathbf{I}+\mathbf{E}_1+\mathbf{E}_2+\mathbf{E}_3$ where $\mathbf{I}$ is the identity matrix and

$$\mathbf{E}_1=\begin{bmatrix}0&0&0\\ \lambda&0&0\\ \lambda^2&0&0\end{bmatrix},\quad \mathbf{E}_2=\begin{bmatrix}0&\lambda&0\\ 0&0&0\\ 0&\lambda&0\end{bmatrix},\quad \mathbf{E}_3=\begin{bmatrix}0&0&\lambda^2\\ 0&0&\lambda\\ 0&0&0\end{bmatrix}.$$

Then $\mathbf{C}_1=\mathbf{I}$ and the formulas for ν_k and $\mathbf{C}_{k+1}^{-1}$ yield $\nu_1=1$ and $\mathbf{C}_2^{-1}=\mathbf{I}-\mathbf{E}_1$. Continuing the computation,

$$\nu_2=\frac{1}{1+\operatorname{tr}\mathbf{C}_2^{-1}\mathbf{E}_2}=\frac{1}{1-\lambda^2}$$

and

$$\begin{aligned}\mathbf{C}_3^{-1}&=\mathbf{C}_2^{-1}-\nu_2\mathbf{C}_2^{-1}\mathbf{E}_2\mathbf{C}_2^{-1}\\ &=(\mathbf{I}-\mathbf{E}_1)-\frac{1}{1-\lambda^2}(\mathbf{I}-\mathbf{E}_1)\mathbf{E}_2(\mathbf{I}-\mathbf{E}_1)\\ &=\begin{bmatrix}\dfrac{1}{1-\lambda^2}&-\dfrac{\lambda}{1-\lambda^2}&0\\ -\dfrac{\lambda}{1-\lambda^2}&\dfrac{1}{1-\lambda^2}&0\\ 0&-\lambda&1\end{bmatrix}.\end{aligned}$$

Finally,

$$\nu_3=\frac{1}{1+\operatorname{tr}\mathbf{C}_3^{-1}\mathbf{E}_3}=\frac{1}{1-\lambda^2}$$

and

$$\begin{aligned}\mathbf{\Lambda}^{-1}&=\mathbf{C}_3^{-1}-\nu_3\mathbf{C}_3^{-1}\mathbf{E}_3\mathbf{C}_3^{-1}\\ &=\mathbf{C}_3^{-1}-\frac{1}{1-\lambda^2}\begin{bmatrix}0&0&0\\ 0&-\lambda^2&\lambda\\ 0&\lambda^3&-\lambda^2\end{bmatrix}\\ &=\frac{1}{1-\lambda^2}\begin{bmatrix}1&-\lambda&0\\ -\lambda&1+\lambda^2&-\lambda\\ 0&-\lambda&1\end{bmatrix}.\end{aligned}$$

Let us now consider some applications involving the Kronecker product of matrices. Suppose $\mathbf{A}$ is an $M\times M$ matrix and $\mathbf{G}=\|g_{mn}\|$ an $N\times N$ matrix. Then the $MN\times MN$ dimensional partitioned matrix

$$\begin{bmatrix}g_{11}\mathbf{A}&g_{12}\mathbf{A}&\cdots&g_{1N}\mathbf{A}\\ g_{21}\mathbf{A}&g_{22}\mathbf{A}&\cdots&g_{2N}\mathbf{A}\\ \cdots&\cdots&\cdots&\cdots\\ g_{N1}\mathbf{A}&g_{N2}\mathbf{A}&\cdots&g_{NN}\mathbf{A}\end{bmatrix}\tag{14}$$

is called the **Kronecker product** or **direct product** of $\mathbf{A}$ and $\mathbf{G}$ and is frequently written as $\mathbf{A}\otimes\mathbf{G}$ [7, page 81]. The algebra of direct products is interesting. For example, if $\mathbf{A}$ and $\mathbf{G}$ are nonsingular, then

$$(\mathbf{A}\otimes\mathbf{G})^{-1}=\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}.\tag{15}$$

(Note that the order of $\mathbf{A}$ and $\mathbf{G}$ is not reversed.) However, the only nontrivial property we really need in our analysis is the formula

$$(\mathbf{A}\otimes\mathbf{G})(\mathbf{B}\otimes\mathbf{H})=\mathbf{AB}\otimes\mathbf{GH}\tag{16}$$

where $\mathbf{A}$ and $\mathbf{B}$ are $M\times M$ matrices and $\mathbf{G}$ and $\mathbf{H}$ are $N\times N$ matrices.

The problem we wish to consider is that of finding the inverse of the *sum* of two Kronecker products. We begin by considering the matrix

$$\mathbf{W}=\mathbf{A}\otimes\mathbf{G}+\mathbf{B}\otimes\mathbf{E} \tag{17}$$

where $\mathbf{E}$ is an $N\times N$ matrix of rank one, and $\mathbf{A}$, $\mathbf{G}$ and $\mathbf{W}$ are nonsingular. Our previous analyses suggest that we search for an inverse in the form

$$\mathbf{W}^{-1}=\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}-\mathbf{T}\otimes\mathbf{G}^{-1}\mathbf{E}\mathbf{G}^{-1} \tag{18}$$

where $\mathbf{T}$ is a matrix to be determined. If we multiply (17) and (18) together and use (15) and (16) we obtain

$$\mathbf{W}\mathbf{W}^{-1}=\mathbf{I}\otimes\mathbf{I}-\mathbf{A}\mathbf{T}\otimes\mathbf{E}\mathbf{G}^{-1}+\mathbf{B}\mathbf{A}^{-1}\otimes\mathbf{E}\mathbf{G}^{-1}-\mathbf{B}\mathbf{T}\otimes\mathbf{E}\mathbf{G}^{-1}\mathbf{E}\mathbf{G}^{-1}.$$

But by our reproducing property (6), $\mathbf{E}\mathbf{G}^{-1}\mathbf{E}\mathbf{G}^{-1}=g\mathbf{E}\mathbf{G}^{-1}$ where $g=\operatorname{tr}\mathbf{E}\mathbf{G}^{-1}$. Hence if $\mathbf{W}\mathbf{W}^{-1}$ is to be the identity matrix, then $(-\mathbf{A}\mathbf{T}+\mathbf{B}\mathbf{A}^{-1}-g\mathbf{B}\mathbf{T})\otimes\mathbf{E}\mathbf{G}^{-1}$ must be the zero matrix. Thus we choose $\mathbf{T}$ to be $\mathbf{T}=(\mathbf{A}+g\mathbf{B})^{-1}\mathbf{B}\mathbf{A}^{-1}$.

Certain special cases are of interest.

EXAMPLE 5. ($\mathbf{G}$ the identity matrix) In this case (17) may be written

$$\mathbf{W}=\mathbf{A}\otimes\mathbf{I}+\mathbf{B}\otimes\mathbf{E} \tag{19}$$

and $\mathbf{T}=[\mathbf{A}+(\operatorname{tr}\mathbf{E})\mathbf{B}]^{-1}\mathbf{B}\mathbf{A}^{-1}$. Thus

$$\mathbf{W}^{-1}=\mathbf{A}^{-1}\otimes\mathbf{I}-[\mathbf{A}+(\operatorname{tr}\mathbf{E})\mathbf{B}]^{-1}\mathbf{B}\mathbf{A}^{-1}\otimes\mathbf{E}. \tag{20}$$

We also observe that $\det\mathbf{W}=(\det\mathbf{A})^{N-1}\det[\mathbf{A}+(\operatorname{tr}\mathbf{E})\mathbf{B}]$; and if $\lambda_1,\dots,\lambda_M$ are the eigenvalues of $\mathbf{A}$, and $\mu_1,\dots,\mu_M$ the eigenvalues of $\mathbf{A}+(\operatorname{tr}\mathbf{E})\mathbf{B}$, then each λ_m, $1\leqslant m\leqslant M$, is an eigenvalue of multiplicity $N-1$ of $\mathbf{W}$ and the μ_m, $1\leqslant m\leqslant M$, are simple eigenvalues of $\mathbf{W}$.

EXAMPLE 6. ($N=1$) If $N=1$ in (19), then $\mathbf{W}=\mathbf{A}+e\mathbf{B}$ where $e=\operatorname{tr}\mathbf{E}\neq 0$. Equation (20) then yields the well-known identity $(\mathbf{A}+e\mathbf{B})^{-1}=\mathbf{A}^{-1}-e(\mathbf{A}+e\mathbf{B})^{-1}\mathbf{B}\mathbf{A}^{-1}$. (Compare this with the result given by the fundamental Lemma.)

Now, returning to (17), suppose $\mathbf{B}=\mathbf{F}$ is also a matrix of rank one. Then $\mathbf{W}=\mathbf{A}\otimes\mathbf{G}+\mathbf{F}\otimes\mathbf{E}$ and by (18) its inverse is

$$\mathbf{W}^{-1}=\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}-(\mathbf{A}+g\mathbf{F})^{-1}\mathbf{F}\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}\mathbf{E}\mathbf{G}^{-1}. \tag{21}$$

We notice now that since $\mathbf{F}$ also is of rank one, the fundamental Lemma may be applied to $\mathbf{A}+g\mathbf{F}$, viz.:

$$(\mathbf{A}+g\mathbf{F})^{-1}=\mathbf{A}^{-1}-\frac{g}{1+ga}\mathbf{A}^{-1}\mathbf{F}\mathbf{A}^{-1}$$

where $a=\operatorname{tr}\mathbf{F}\mathbf{A}^{-1}$. Substituting this result in (21) and simplifying yields the elegant formula:

$$(\mathbf{A}\otimes\mathbf{G}+\mathbf{F}\otimes\mathbf{E})^{-1}=\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}-\frac{1}{1+ga}\mathbf{A}^{-1}\mathbf{F}\mathbf{A}^{-1}\otimes\mathbf{G}^{-1}\mathbf{E}\mathbf{G}^{-1}.$$

References

[1] G. Dahlquist and Å. Björck, Numerical Methods, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[2] P. S. Dwyer and F. V. Waugh, On errors in matrix inversion, J. Amer. Statist. Assoc., 48 (1953) 289–319.

[3] D. K. Faddeev and V. N. Faddeeva, Computational Methods of Linear Algebra, Freeman, San Francisco, 1963.

[4] A. Gelb, Applied Optimal Estimation, M.I.T. Press, Cambridge, 1974.

[5] P. R. Halmos, Finite-Dimensional Vector Spaces, Van Nostrand, Princeton, 1958.

[6] D. M. Leskiw and K. S. Miller, A comparison of some Kalman estimators, IEEE Trans., IT-25 (1979) 491–495.

[7] C. C. Mac Duffee, The Theory of Matrices, Chelsea, New York, 1946.

[8] K. S. Miller, Vector Stochastic Processes, Krieger, Huntington, NY, 1980.