3.7 MATRIX INVERSION

If α is a nonzero scalar, then for each number β the equation $\alpha x = \beta$ has a unique solution given by $x = \alpha^{-1}\beta$. To *prove* that $\alpha^{-1}\beta$ is a solution, write

$$\alpha(\alpha^{-1}\beta) = (\alpha\alpha^{-1})\beta = (1)\beta = \beta. \tag{3.7.1}$$

Uniqueness follows because if x_1 and x_2 are two solutions, then

$$\begin{aligned} \alpha x_1 = \beta = \alpha x_2 \quad &\Longrightarrow \quad \alpha^{-1}(\alpha x_1) = \alpha^{-1}(\alpha x_2) \\ &\Longrightarrow \quad (\alpha^{-1}\alpha)x_1 = (\alpha^{-1}\alpha)x_2 \\ &\Longrightarrow \quad (1)x_1 = (1)x_2 \quad \Longrightarrow \quad x_1 = x_2. \end{aligned} \tag{3.7.2}$$

These observations seem pedantic, but they are important in order to see how to make the transition from scalar equations to matrix equations. In particular, these arguments show that in addition to associativity, the properties

$$\alpha\alpha^{-1} = 1 \quad \text{and} \quad \alpha^{-1}\alpha = 1 \tag{3.7.3}$$

are the key ingredients, so if we want to solve matrix equations in the same fashion as we solve scalar equations, then a matrix analogue of (3.7.3) is needed.

Matrix Inversion

For a given square matrix $\mathbf{A}_{n\times n}$, the matrix $\mathbf{B}_{n\times n}$ that satisfies the conditions

$$\mathbf{AB} = \mathbf{I}_n \quad \text{and} \quad \mathbf{BA} = \mathbf{I}_n$$

is called the ***inverse*** of $\mathbf{A}$ and is denoted by $\mathbf{B} = \mathbf{A}^{-1}$. Not all square matrices are invertible—the zero matrix is a trivial example, but there are also many nonzero matrices that are not invertible. An invertible matrix is said to be ***nonsingular,*** and a square matrix with no inverse is called a ***singular matrix.***

Notice that matrix inversion is defined for square matrices only—the condition $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A}$ rules out inverses of nonsquare matrices.

Example 3.7.1

If

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where} \quad \delta = ad - bc \neq 0,$$

then

$$\mathbf{A}^{-1} = \frac{1}{\delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

because it can be verified that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_2$.

Although not all matrices are invertible, *when an inverse exists, it is unique.* To see this, suppose that $\mathbf{X}_1$ and $\mathbf{X}_2$ are both inverses for a nonsingular matrix $\mathbf{A}$. Then

$$\mathbf{X}_1 = \mathbf{X}_1\mathbf{I} = \mathbf{X}_1(\mathbf{A}\mathbf{X}_2) = (\mathbf{X}_1\mathbf{A})\mathbf{X}_2 = \mathbf{I}\mathbf{X}_2 = \mathbf{X}_2,$$

which implies that only one inverse is possible.

Since matrix inversion was defined analogously to scalar inversion, and since matrix multiplication is associative, exactly the same reasoning used in (3.7.1) and (3.7.2) can be applied to a matrix equation $\mathbf{AX} = \mathbf{B}$, so we have the following statements.

Matrix Equations

- If $\mathbf{A}$ is a nonsingular matrix, then there is a unique solution for $\mathbf{X}$ in the matrix equation $\mathbf{A}_{n\times n}\mathbf{X}_{n\times p} = \mathbf{B}_{n\times p}$, and the solution is

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}. \tag{3.7.4}$$

- A system of n linear equations in n unknowns can be written as a single matrix equation $\mathbf{A}_{n\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{n\times 1}$ (see p. 99), so it follows from (3.7.4) that when $\mathbf{A}$ is nonsingular, the system has a unique solution given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

However, it must be stressed that the representation of the solution as $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ is mostly a notational or theoretical convenience. *In practice, a nonsingular system* $\mathbf{Ax} = \mathbf{b}$ *is almost never solved by first computing* $\mathbf{A}^{-1}$ *and then the product* $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. The reason will be apparent when we learn how much work is involved in computing $\mathbf{A}^{-1}$.

Since not all square matrices are invertible, methods are needed to distinguish between nonsingular and singular matrices. There is a variety of ways to describe the class of nonsingular matrices, but those listed below are among the most important.

Existence of an Inverse

For an $n \times n$ matrix $\mathbf{A}$, the following statements are equivalent.

- $\mathbf{A}^{-1}$ exists ($\mathbf{A}$ is nonsingular). (3.7.5)
- $rank\,(\mathbf{A}) = n$. (3.7.6)
- $\mathbf{A} \xrightarrow{\text{Gauss–Jordan}} \mathbf{I}$. (3.7.7)
- $\mathbf{Ax} = \mathbf{0}$ implies that $\mathbf{x} = \mathbf{0}$. (3.7.8)

Proof. The fact that (3.7.6) $\Longleftrightarrow$ (3.7.7) is a direct consequence of the definition of rank, and (3.7.6) $\Longleftrightarrow$ (3.7.8) was established in §2.4. Consequently, statements (3.7.6), (3.7.7), and (3.7.8) are equivalent, so if we establish that (3.7.5) $\Longleftrightarrow$ (3.7.6), then the proof will be complete.

Proof of (3.7.5) $\Longrightarrow$ (3.7.6). Begin by observing that (3.5.5) guarantees that a matrix $\mathbf{X} = [\mathbf{X}_{*1} \,|\, \mathbf{X}_{*2} \,|\, \cdots \,|\, \mathbf{X}_{*n}]$ satisfies the equation $\mathbf{AX} = \mathbf{I}$ if and only if $\mathbf{X}_{*j}$ is a solution of the linear system $\mathbf{Ax} = \mathbf{I}_{*j}$. If $\mathbf{A}$ is nonsingular, then we know from (3.7.4) that there exists a unique solution to $\mathbf{AX} = \mathbf{I}$, and hence each linear system $\mathbf{Ax} = \mathbf{I}_{*j}$ has a unique solution. But in §2.5 we learned that a linear system has a unique solution if and only if the rank of the coefficient matrix equals the number of unknowns, so $rank\,(\mathbf{A}) = n$.

Proof of (3.7.6) $\Longrightarrow$ (3.7.5). If $rank\,(\mathbf{A}) = n$, then (2.3.4) insures that each system $\mathbf{Ax} = \mathbf{I}_{*j}$ is consistent because $rank[\mathbf{A} \,|\, \mathbf{I}_{*j}] = n = rank\,(\mathbf{A})$. Furthermore, the results of §2.5 guarantee that each system $\mathbf{Ax} = \mathbf{I}_{*j}$ has a unique solution, and hence there is a unique solution to the matrix equation $\mathbf{AX} = \mathbf{I}$. We would like to say that $\mathbf{X} = \mathbf{A}^{-1}$, but we cannot jump to this conclusion without first arguing that $\mathbf{XA} = \mathbf{I}$. Suppose this is not true—i.e., suppose that $\mathbf{XA} - \mathbf{I} \neq \mathbf{0}$. Since

$$\mathbf{A}(\mathbf{XA} - \mathbf{I}) = (\mathbf{AX})\mathbf{A} - \mathbf{A} = \mathbf{IA} - \mathbf{A} = \mathbf{0},$$

it follows from (3.5.5) that any nonzero column of $\mathbf{XA} - \mathbf{I}$ is a nontrivial solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$. But this is a contradiction of the fact that (3.7.6) $\Longleftrightarrow$ (3.7.8). Therefore, the supposition that $\mathbf{XA} - \mathbf{I} \neq \mathbf{0}$ must be false, and thus $\mathbf{AX} = \mathbf{I} = \mathbf{XA}$, which means $\mathbf{A}$ is nonsingular. ∎

The definition of matrix inversion says that in order to compute $\mathbf{A}^{-1}$, it is necessary to solve *both* of the matrix equations $\mathbf{AX} = \mathbf{I}$ and $\mathbf{XA} = \mathbf{I}$. These two equations are necessary to rule out the possibility of nonsquare inverses. But when only square matrices are involved, then any one of the two equations will suffice—the following example elaborates.

Example 3.7.2

Problem: If $\mathbf{A}$ and $\mathbf{X}$ are *square* matrices, explain why

$$\mathbf{AX} = \mathbf{I} \quad \Longrightarrow \quad \mathbf{XA} = \mathbf{I}. \tag{3.7.9}$$

In other words, if $\mathbf{A}$ and $\mathbf{X}$ are square and $\mathbf{AX} = \mathbf{I}$, then $\mathbf{X} = \mathbf{A}^{-1}$.

Solution: Notice first that $\mathbf{AX} = \mathbf{I}$ implies $\mathbf{X}$ is nonsingular because if $\mathbf{X}$ is singular, then, by (3.7.8), there is a column vector $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{Xx} = \mathbf{0}$, which is contrary to the fact that $\mathbf{x} = \mathbf{Ix} = \mathbf{AXx} = \mathbf{0}$. Now that we know $\mathbf{X}^{-1}$ exists, we can establish (3.7.9) by writing

$$\mathbf{AX} = \mathbf{I} \quad \Longrightarrow \quad \mathbf{AXX}^{-1} = \mathbf{X}^{-1} \quad \Longrightarrow \quad \mathbf{A} = \mathbf{X}^{-1} \quad \Longrightarrow \quad \mathbf{XA} = \mathbf{I}.$$

Caution! The argument above is not valid for nonsquare matrices. When $m \neq n$, it's possible that $\mathbf{A}_{m\times n}\mathbf{X}_{n\times m} = \mathbf{I}_m$, but $\mathbf{XA} \neq \mathbf{I}_n$.

Although we usually try to avoid computing the inverse of a matrix, there are times when an inverse must be found. To construct an algorithm that will yield $\mathbf{A}^{-1}$ when $\mathbf{A}_{n\times n}$ is nonsingular, recall from Example 3.7.2 that determining $\mathbf{A}^{-1}$ is equivalent to solving the single matrix equation $\mathbf{AX} = \mathbf{I}$, and due to (3.5.5), this in turn is equivalent to solving the n linear systems defined by

$$\mathbf{Ax} = \mathbf{I}_{*j} \quad \text{for} \quad j = 1, 2, \ldots, n. \tag{3.7.10}$$

In other words, if $\mathbf{X}_{*1}, \mathbf{X}_{*2}, \ldots, \mathbf{X}_{*n}$ are the respective solutions to (3.7.10), then $\mathbf{X} = [\mathbf{X}_{*1} \,|\, \mathbf{X}_{*2} \,|\, \cdots \,|\, \mathbf{X}_{*n}]$ solves the equation $\mathbf{AX} = \mathbf{I}$, and hence $\mathbf{X} = \mathbf{A}^{-1}$. If $\mathbf{A}$ is nonsingular, then we know from (3.7.7) that the Gauss–Jordan method reduces the augmented matrix $[\mathbf{A} \,|\, \mathbf{I}_{*j}]$ to $[\mathbf{I} \,|\, \mathbf{X}_{*j}]$, and the results of §1.3 insure that $\mathbf{X}_{*j}$ is the unique solution to $\mathbf{Ax} = \mathbf{I}_{*j}$. That is,

$$[\mathbf{A} \,|\, \mathbf{I}_{*j}] \xrightarrow{\text{Gauss–Jordan}} \left[\mathbf{I} \,\middle|\, [\mathbf{A}^{-1}]_{*j}\right].$$

But rather than solving each system $\mathbf{Ax} = \mathbf{I}_{*j}$ separately, we can solve them simultaneously by taking advantage of the fact that they all have the same coefficient matrix. In other words, applying the Gauss–Jordan method to the larger augmented array $[\mathbf{A} \,|\, \mathbf{I}_{*1} \,|\, \mathbf{I}_{*2} \,|\, \cdots \,|\, \mathbf{I}_{*n}]$ produces

$$[\mathbf{A} \,|\, \mathbf{I}_{*1} \,|\, \mathbf{I}_{*2} \,|\, \cdots \,|\, \mathbf{I}_{*n}] \xrightarrow{\text{Gauss–Jordan}} \left[\mathbf{I} \,\middle|\, [\mathbf{A}^{-1}]_{*1} \,\middle|\, [\mathbf{A}^{-1}]_{*2} \,\middle|\, \cdots \,\middle|\, [\mathbf{A}^{-1}]_{*n}\right],$$

or more compactly,

$$[\mathbf{A} \,|\, \mathbf{I}] \xrightarrow{\text{Gauss–Jordan}} [\mathbf{I} \,|\, \mathbf{A}^{-1}]. \tag{3.7.11}$$

What happens if we try to invert a singular matrix using this procedure? The fact that (3.7.5) $\Longleftrightarrow$ (3.7.6) $\Longleftrightarrow$ (3.7.7) guarantees that a singular matrix $\mathbf{A}$ cannot be reduced to $\mathbf{I}$ by Gauss–Jordan elimination because a zero row will have to emerge in the left-hand side of the augmented array at some point during the process. This means that we do not need to know at the outset whether $\mathbf{A}$ is nonsingular or singular—it becomes self-evident depending on whether or not the reduction (3.7.11) can be completed. A summary is given below.

Computing an Inverse

Gauss–Jordan elimination can be used to invert $\mathbf{A}$ by the reduction

$$[\mathbf{A} \,|\, \mathbf{I}] \xrightarrow{\text{Gauss–Jordan}} [\mathbf{I} \,|\, \mathbf{A}^{-1}]. \tag{3.7.12}$$

The only way for this reduction to fail is for a row of zeros to emerge in the left-hand side of the augmented array, and this occurs if and only if $\mathbf{A}$ is a singular matrix. A different (and somewhat more practical) algorithm is given Example 3.10.3 on p. 148.

Although they are not included in the simple examples of this section, you are reminded that the pivoting and scaling strategies presented in §1.5 need to be incorporated, and the effects of ill-conditioning discussed in §1.6 must be considered whenever matrix inverses are computed using floating-point arithmetic. However, practical applications rarely require an inverse to be computed.

Example 3.7.3

Problem: If possible, find the inverse of $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

Solution:

$$[\mathbf{A}\,|\,\mathbf{I}] = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right)$$

Therefore, the matrix is nonsingular, and $\mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$. If we wish to check this answer, we need only check that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. If this holds, then the result of Example 3.7.2 insures that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ will automatically be true.

Earlier in this section it was stated that one almost never solves a nonsingular linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ by first computing $\mathbf{A}^{-1}$ and then the product $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. To appreciate why this is true, pay attention to how much effort is required to perform one matrix inversion.

Operation Counts for Inversion

Computing $\mathbf{A}^{-1}_{n\times n}$ by reducing $[\mathbf{A}|\mathbf{I}]$ with Gauss–Jordan requires

- n^3 multiplications/divisions,
- $n^3 - 2n^2 + n$ additions/subtractions.

Interestingly, if Gaussian elimination with a back substitution process is applied to $[\mathbf{A}|\mathbf{I}]$ instead of the Gauss–Jordan technique, then exactly the same operation count can be obtained. Although Gaussian elimination with back substitution is more efficient than the Gauss–Jordan method for solving a single linear system, the two procedures are essentially equivalent for inversion.

Solving a nonsingular system $\mathbf{Ax} = \mathbf{b}$ by first computing $\mathbf{A}^{-1}$ and then forming the product $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ requires $n^3 + n^2$ multiplications/divisions and $n^3 - n^2$ additions/subtractions. Recall from §1.5 that Gaussian elimination with back substitution requires only about $n^3/3$ multiplications/divisions and about $n^3/3$ additions/subtractions. In other words, *using* $\mathbf{A}^{-1}$ *to solve a nonsingular system* $\mathbf{Ax} = \mathbf{b}$ *requires about three times the effort as does Gaussian elimination with back substitution.*

To put things in perspective, consider standard matrix multiplication between two $n \times n$ matrices. It is not difficult to verify that n^3 multiplications and $n^3 - n^2$ additions are required. Remarkably, it takes almost exactly as much effort to perform one matrix multiplication as to perform one matrix inversion. This fact always seems to be counter to a novice's intuition—it "feels" like matrix inversion should be a more difficult task than matrix multiplication, but this is not the case.

The remainder of this section is devoted to a discussion of some of the important properties of matrix inversion. We begin with the four basic facts listed below.

Properties of Matrix Inversion

For nonsingular matrices $\mathbf{A}$ and $\mathbf{B}$, the following properties hold.

- $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$. (3.7.13)
- The product $\mathbf{AB}$ is also nonsingular. (3.7.14)
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (the reverse order law for inversion). (3.7.15)
- $\left(\mathbf{A}^{-1}\right)^T = \left(\mathbf{A}^T\right)^{-1}$ and $\left(\mathbf{A}^{-1}\right)^* = (\mathbf{A}^*)^{-1}$. (3.7.16)

Proof. Property (3.7.13) follows directly from the definition of inversion. To prove (3.7.14) and (3.7.15), let $\mathbf{X} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ and verify that $(\mathbf{AB})\mathbf{X} = \mathbf{I}$ by writing

$$(\mathbf{AB})\mathbf{X} = (\mathbf{AB})\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{A}(\mathbf{I})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}.$$

According to the discussion in Example 3.7.2, we are now guaranteed that $\mathbf{X}(\mathbf{AB}) = \mathbf{I}$, and we need not bother to verify it. To prove property (3.7.16), let $\mathbf{X} = \left(\mathbf{A}^{-1}\right)^T$ and verify that $\mathbf{A}^T\mathbf{X} = \mathbf{I}$. Make use of the reverse order law for transposition to write

$$\mathbf{A}^T\mathbf{X} = \mathbf{A}^T\left(\mathbf{A}^{-1}\right)^T = \left(\mathbf{A}^{-1}\mathbf{A}\right)^T = \mathbf{I}^T = \mathbf{I}.$$

Therefore, $\left(\mathbf{A}^T\right)^{-1} = \mathbf{X} = \left(\mathbf{A}^{-1}\right)^T$. The proof of the conjugate transpose case is similar. ∎

In general the product of two rank-r matrices does not necessarily have to produce another matrix of rank r. For example,

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}$$

each has rank 1, but the product $\mathbf{AB} = \mathbf{0}$ has rank 0. However, we saw in (3.7.14) that the product of two invertible matrices is again invertible. That is, if $rank\,(A_{n\times n}) = n$ and $rank\,(B_{n\times n}) = n$, then $rank\,(\mathbf{AB}) = n$. This generalizes to any number of matrices.

Products of Nonsingular Matrices Are Nonsingular

If $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ are each $n \times n$ nonsingular matrices, then the product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$ is also nonsingular, and its inverse is given by the reverse order law. That is,

$$(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\cdots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.$$

Proof. Apply (3.7.14) and (3.7.15) inductively. For example, when $k = 3$ you can write

$$(\mathbf{A}_1\{\mathbf{A}_2\mathbf{A}_3\})^{-1} = \{\mathbf{A}_2\mathbf{A}_3\}^{-1}\mathbf{A}_1^{-1} = \mathbf{A}_3^{-1}\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}. \quad \blacksquare$$

Exercises for section 3.7

3.7.1. When possible, find the inverse of each of the following matrices. Check your answer by using matrix multiplication.

$$\text{(a)}\ \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{(b)}\ \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \text{(c)}\ \begin{pmatrix} 4 & -8 & 5 \\ 4 & -7 & 4 \\ 3 & -4 & 2 \end{pmatrix}$$

$$\text{(d)}\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \text{(e)}\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

3.7.2. Find the matrix $\mathbf{X}$ such that $\mathbf{X} = \mathbf{AX} + \mathbf{B}$, where

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{pmatrix}.$$

3.7.3. For a square matrix $\mathbf{A}$, explain why each of the following statements must be true.

(a) If $\mathbf{A}$ contains a zero row or a zero column, then $\mathbf{A}$ is singular.

(b) If $\mathbf{A}$ contains two identical rows or two identical columns, then $\mathbf{A}$ is singular.

(c) If one row (or column) is a multiple of another row (or column), then $\mathbf{A}$ must be singular.

3.7.4. Answer each of the following questions.

(a) Under what conditions is a diagonal matrix nonsingular? Describe the structure of the inverse of a diagonal matrix.

(b) Under what conditions is a triangular matrix nonsingular? Describe the structure of the inverse of a triangular matrix.

3.7.5. If $\mathbf{A}$ is nonsingular and symmetric, prove that $\mathbf{A}^{-1}$ is symmetric.

3.7.6. If $\mathbf{A}$ is a square matrix such that $\mathbf{I}-\mathbf{A}$ is nonsingular, prove that

$$\mathbf{A}(\mathbf{I}-\mathbf{A})^{-1} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{A}.$$

3.7.7. Prove that if $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times m$ such that $\mathbf{AB} = \mathbf{I}_m$ and $\mathbf{BA} = \mathbf{I}_n$, then $m = n$.

3.7.8. If $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$ are each nonsingular, prove that

$$\mathbf{A}(\mathbf{A}+\mathbf{B})^{-1}\mathbf{B} = \mathbf{B}(\mathbf{A}+\mathbf{B})^{-1}\mathbf{A} = \left(\mathbf{A}^{-1}+\mathbf{B}^{-1}\right)^{-1}.$$

3.7.9. Let $\mathbf{S}$ be a skew-symmetric matrix with real entries.

(a) Prove that $\mathbf{I}-\mathbf{S}$ is nonsingular. **Hint:** $\mathbf{x}^T\mathbf{x} = 0 \implies \mathbf{x} = \mathbf{0}$.

(b) If $\mathbf{A} = (\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1}$, show that $\mathbf{A}^{-1} = \mathbf{A}^T$.

3.7.10. For matrices $\mathbf{A}_{r\times r}$, $\mathbf{B}_{s\times s}$, and $\mathbf{C}_{r\times s}$ such that $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, verify that each of the following is true.

(a) $$\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{pmatrix}$$

(b) $$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1} \\ \mathbf{0} & \mathbf{B}^{-1} \end{pmatrix}$$

3.7.11. Consider the block matrix $\begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{C}_{r\times s} \\ \mathbf{R}_{s\times r} & \mathbf{B}_{s\times s} \end{pmatrix}$. When the indicated inverses exist, the matrices defined by

$$\mathbf{S} = \mathbf{B} - \mathbf{R}\mathbf{A}^{-1}\mathbf{C} \quad \text{and} \quad \mathbf{T} = \mathbf{A} - \mathbf{C}\mathbf{B}^{-1}\mathbf{R}$$

are called the ***Schur complements***[20] of $\mathbf{A}$ and $\mathbf{B}$, respectively.

(a) If $\mathbf{A}$ and $\mathbf{S}$ are both nonsingular, verify that

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{R} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{C}\mathbf{S}^{-1}\mathbf{R}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{R}\mathbf{A}^{-1} & \mathbf{S}^{-1} \end{pmatrix}.$$

(b) If $\mathbf{B}$ and $\mathbf{T}$ are nonsingular, verify that

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{R} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{T}^{-1} & -\mathbf{T}^{-1}\mathbf{C}\mathbf{B}^{-1} \\ -\mathbf{B}^{-1}\mathbf{R}\mathbf{T}^{-1} & \mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{R}\mathbf{T}^{-1}\mathbf{C}\mathbf{B}^{-1} \end{pmatrix}.$$

3.7.12. Suppose that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are $n \times n$ matrices such that $\mathbf{A}\mathbf{B}^T$ and $\mathbf{C}\mathbf{D}^T$ are each symmetric and $\mathbf{A}\mathbf{D}^T - \mathbf{B}\mathbf{C}^T = \mathbf{I}$. Prove that

$$\mathbf{A}^T\mathbf{D} - \mathbf{C}^T\mathbf{B} = \mathbf{I}.$$

---

[20] This is named in honor of the German mathematician Issai Schur (1875–1941), who first studied matrices of this type. Schur was a student and collaborator of Ferdinand Georg Frobenius (p. 662). Schur and Frobenius were among the first to study matrix theory as a discipline unto itself, and each made great contributions to the subject. It was Emilie V. Haynsworth (1916–1987)—a mathematical granddaughter of Schur—who introduced the phrase "Schur complement" and developed several important aspects of the concept.

Solutions for exercises in section 3. 7

3.7.1. (a) $\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$ (b) Singular (c) $\begin{pmatrix} 2 & -4 & 3 \\ 4 & -7 & 4 \\ 5 & -8 & 4 \end{pmatrix}$ (d) Singular

(e) $\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$

3.7.2. Write the equation as $(\mathbf{I}-\mathbf{A})\mathbf{X}=\mathbf{B}$ and compute

$$\mathbf{X}=(\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}=\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{pmatrix}=\begin{pmatrix} 2 & 4 \\ -1 & -2 \\ 3 & 3 \end{pmatrix}.$$

3.7.3. In each case, the given information implies that $rank(\mathbf{A})<n$—see the solution for Exercise 2.1.3.

3.7.4. (a) If $\mathbf{D}$ is diagonal, then $\mathbf{D}^{-1}$ exists if and only if each $d_{ii}\neq 0$, in which case

$$\begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}^{-1}=\begin{pmatrix} 1/d_{11} & 0 & \cdots & 0 \\ 0 & 1/d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_{nn} \end{pmatrix}.$$

(b) If $\mathbf{T}$ is triangular, then $\mathbf{T}^{-1}$ exists if and only if each $t_{ii}\neq 0$. If $\mathbf{T}$ is upper (lower) triangular, then $\mathbf{T}^{-1}$ is also upper (lower) triangular with $[\mathbf{T}^{-1}]_{ii}=1/t_{ii}$.

3.7.5. $\left(\mathbf{A}^{-1}\right)^T=\left(\mathbf{A}^T\right)^{-1}=\mathbf{A}^{-1}$.

3.7.6. Start with $\mathbf{A}(\mathbf{I}-\mathbf{A})=(\mathbf{I}-\mathbf{A})\mathbf{A}$ and apply $(\mathbf{I}-\mathbf{A})^{-1}$ to both sides, first on one side and then on the other.

3.7.7. Use the result of Example 3.6.5 that says that $trace(\mathbf{AB})=trace(\mathbf{BA})$ to write

$$m=trace(\mathbf{I}_m)=trace(\mathbf{AB})=trace(\mathbf{BA})=trace(\mathbf{I}_n)=n.$$

3.7.8. Use the reverse order law for inversion to write

$$\left[\mathbf{A}(\mathbf{A}+\mathbf{B})^{-1}\mathbf{B}\right]^{-\mathbf{1}}=\mathbf{B}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{A}^{-1}=\mathbf{B}^{-1}+\mathbf{A}^{-1}$$

and

$$\left[\mathbf{B}(\mathbf{A}+\mathbf{B})^{-1}\mathbf{A}\right]^{-\mathbf{1}}=\mathbf{A}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{B}^{-1}=\mathbf{B}^{-1}+\mathbf{A}^{-1}.$$

3.7.9. (a) $(\mathbf{I}-\mathbf{S})\mathbf{x}=\mathbf{0} \implies \mathbf{x}^T(\mathbf{I}-\mathbf{S})\mathbf{x}=0 \implies \mathbf{x}^T\mathbf{x}=\mathbf{x}^T\mathbf{S}\mathbf{x}$. Taking transposes on both sides yields $\mathbf{x}^T\mathbf{x}=-\mathbf{x}^T\mathbf{S}\mathbf{x}$, so that $\mathbf{x}^T\mathbf{x}=0$, and thus $\mathbf{x}=\mathbf{0}$

(recall Exercise 3.6.12). The conclusion follows from property (3.7.8).
(b) First notice that Exercise 3.7.6 implies that $\mathbf{A} = (\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1} = (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})$. By using the reverse order laws, transposing both sides yields exactly the same thing as inverting both sides.

3.7.10. Use block multiplication to verify that the product of the matrix with its inverse is the identity matrix.

3.7.11. Use block multiplication to verify that the product of the matrix with its inverse is the identity matrix.

3.7.12. Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ and $\mathbf{X} = \begin{pmatrix} \mathbf{D}^T & -\mathbf{B}^T \\ -\mathbf{C}^T & \mathbf{A}^T \end{pmatrix}$. The hypothesis implies that $\mathbf{MX} = \mathbf{I}$, and hence (from the discussion in Example 3.7.2) it must also be true that $\mathbf{XM} = \mathbf{I}$, from which the conclusion follows. **Note:** This problem appeared on a past Putnam Exam—a national mathematics competition for undergraduate students that is considered to be quite challenging. This means that you can be proud of yourself if you solved it before looking at this solution.

Solutions for exercises in section 3. 8

3.8.1. (a) $\mathbf{B}^{-1} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 4 & -2 \end{pmatrix}$

(b) Let $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ and $\mathbf{d}^T = (0 \quad 2 \quad 1)$ to obtain $\mathbf{C}^{-1} = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 3 & -1 \\ -1 & -4 & 2 \end{pmatrix}$

3.8.2. $\mathbf{A}_{*j}$ needs to be removed, and $\mathbf{b}$ needs to be inserted in its place. This is accomplished by writing $\mathbf{B} = \mathbf{A} + (\mathbf{b} - \mathbf{A}_{*j})\mathbf{e}_j^T$. Applying the Sherman–Morrison formula with $\mathbf{c} = \mathbf{b} - \mathbf{A}_{*j}$ and $\mathbf{d}^T = \mathbf{e}_j^T$ yields

$$\begin{aligned}\mathbf{B}^{-1} &= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}(\mathbf{b}-\mathbf{A}_{*j})\mathbf{e}_j^T\mathbf{A}^{-1}}{1+\mathbf{e}_j^T\mathbf{A}^{-1}(\mathbf{b}-\mathbf{A}_{*j})} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{e}_j^T\mathbf{A}^{-1} - \mathbf{e}_j\mathbf{e}_j^T\mathbf{A}^{-1}}{1+\mathbf{e}_j^T\mathbf{A}^{-1}\mathbf{b} - \mathbf{e}_j^T\mathbf{e}_j} \\ &= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{b}[\mathbf{A}^{-1}]_{j*} - \mathbf{e}_j[\mathbf{A}^{-1}]_{j*}}{[\mathbf{A}^{-1}]_{j*}\mathbf{b}} = \mathbf{A}^{-1} - \frac{\big(\mathbf{A}^{-1}\mathbf{b} - \mathbf{e}_j\big)[\mathbf{A}^{-1}]_{j*}}{[\mathbf{A}^{-1}]_{j*}\mathbf{b}}.\end{aligned}$$

3.8.3. Use the Sherman–Morrison formula to write

$$\begin{aligned}\mathbf{z} &= (\mathbf{A}+\mathbf{c}\mathbf{d}^T)^{-1}\mathbf{b} = \left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^T\mathbf{A}^{-1}}{1+\mathbf{d}^T\mathbf{A}^{-1}\mathbf{c}}\right)\mathbf{b} = \mathbf{A}^{-1}\mathbf{b} - \frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^T\mathbf{A}^{-1}\mathbf{b}}{1+\mathbf{d}^T\mathbf{A}^{-1}\mathbf{c}} \\ &= \mathbf{x} - \frac{\mathbf{y}\mathbf{d}^T\mathbf{x}}{1+\mathbf{d}^T\mathbf{y}}.\end{aligned}$$

3.8.4. (a) For a nonsingular matrix $\mathbf{A}$, the Sherman–Morrison formula guarantees that $\mathbf{A} + \alpha\mathbf{e}_i\mathbf{e}_j^T$ is also nonsingular when $1 + \alpha\left[\mathbf{A}^{-1}\right]_{ji} \neq 0$, and this certainly will be true if α is sufficiently small.