Chapter 11

Matrix inverses

In this chapter we introduce the concept of matrix inverse. We show how matrix inverses can be used to solve linear equations, and how they can be computed using the QR factorization.

11.1 Left and right inverses

Recall that for a number a, its (multiplicative) inverse is the number x for which $xa = 1$, which we usually denote as $x = 1/a$ or (less frequently) $x = a^{-1}$. The inverse x exists provided a is nonzero. For matrices the concept of inverse is more complicated than for scalars; in the general case, we need to distinguish between left and right inverses. We start with the left inverse.

Left inverse. A matrix X that satisfies

$$XA = I$$

is called a *left inverse* of A. The matrix A is said to be *left-invertible* if a left inverse exists. Note that if A has size $m \times n$, a left inverse X will have size $n \times m$, the same dimensions as A^T.

Examples.

- If A is a number (*i.e.*, a 1×1 matrix), then a left inverse X is the same as the inverse of the number. In this case, A is left-invertible whenever A is nonzero, and it has only one left inverse.
- Any nonzero n-vector a, considered as an $n \times 1$ matrix, is left-invertible. For any index i with $a_i \neq 0$, the row n-vector $x = (1/a_i)e_i^T$ satisfies $xa = 1$.
- The matrix

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}$$

has two different left inverses:

$$B = \frac{1}{9}\begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \qquad C = \frac{1}{2}\begin{bmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{bmatrix}.$$

This can be verified by checking that $BA = CA = I$. The example illustrates that a left-invertible matrix can have more than one left inverse. (In fact, if it has more than one left inverse, then it has infinitely many; see exercise 11.1.)

- A matrix A with orthonormal columns satisfies $A^T A = I$, so it is left-invertible; its transpose A^T is a left inverse.

Left-invertibility and column independence. If A has a left inverse C then the columns of A are linearly independent. To see this, suppose that $Ax = 0$. Multiplying on the left by a left inverse C, we get

$$0 = C(Ax) = (CA)x = Ix = x,$$

which shows that the only linear combination of the columns of A that is 0 is the one with all coefficients zero.

We will see below that the converse is also true; a matrix has a left inverse if and only if its columns are linearly independent. So the generalization of 'a number has an inverse if and only if it is nonzero' is 'a matrix has a left inverse if and only if its columns are linearly independent'.

Dimensions of left inverses. Suppose the $m \times n$ matrix A is wide, *i.e.*, $m < n$. By the independence-dimension inequality, its columns are linearly dependent, and therefore it is not left-invertible. Only square or tall matrices can be left-invertible.

Solving linear equations with a left inverse. Suppose that $Ax = b$, where A is an $m \times n$ matrix and x is an n-vector. If C is a left inverse of A, we have

$$Cb = C(Ax) = (CA)x = Ix = x,$$

which means that $x = Cb$ is a solution of the set of linear equations. The columns of A are linearly independent (since it has a left inverse), so there is only one solution of the linear equations $Ax = b$; in other words, $x = Cb$ is *the* solution of $Ax = b$.

Now suppose there is no x that satisfies the linear equations $Ax = b$, and let C be a left inverse of A. Then $x = Cb$ does not satisfy $Ax = b$, since no vector satisfies this equation by assumption. This gives a way to check if the linear equations $Ax = b$ have a solution, and to find one when there is one, provided we have a left inverse of A. We simply test whether $A(Cb) = b$. If this holds, then we have found a solution of the linear equations; if it does not, then we can conclude that there is no solution of $Ax = b$.

In summary, a left inverse can be used to determine whether or not a solution of an over-determined set of linear equations exists, and when it does, find the unique solution.

Right inverse. Now we turn to the closely related concept of right inverse. A matrix X that satisfies

$$AX = I$$

is called a *right inverse* of A. The matrix A is *right-invertible* if a right inverse exists. Any right inverse has the same dimensions as A^T.

Left and right inverse of matrix transpose. If A has a right inverse B, then B^T is a left inverse of A^T, since $B^T A^T = (AB)^T = I$. If A has a left inverse C, then C^T is a right inverse of A^T, since $A^T C^T = (CA)^T = I$. This observation allows us to map all the results for left-invertibility given above to similar results for right-invertibility. Some examples are given below.

- A matrix is right-invertible if and only if its rows are linearly independent.
- A tall matrix cannot have a right inverse. Only square or wide matrices can be right-invertible.

Solving linear equations with a right inverse. Consider the set of m linear equations in n variables $Ax = b$. Suppose A is right-invertible, with right inverse B. This implies that A is square or wide, so the linear equations $Ax = b$ are square or under-determined.

Then for *any* m-vector b, the n-vector $x = Bb$ satisfies the equation $Ax = b$. To see this, we note that

$$Ax = A(Bb) = (AB)b = Ib = b.$$

We can conclude that if A is right-invertible, then the linear equations $Ax = b$ can be solved for *any* vector b. Indeed, $x = Bb$ is a solution. (There can be other solutions of $Ax = b$; the solution $x = Bb$ is simply one of them.)

In summary, a right inverse can be used to find *a* solution of a square or under-determined set of linear equations, for any vector b.

Examples. Consider the matrix appearing in the example above on page 199,

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}$$

and the two left inverses

$$B = \frac{1}{9}\begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \qquad C = \frac{1}{2}\begin{bmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{bmatrix}.$$

- The over-determined linear equations $Ax = (1, -2, 0)$ have the unique solution $x = (1, -1)$, which can be obtained from *either* left inverse:

  $$x = B(1, -2, 0) = C(1, -2, 0).$$

- The over-determined linear equations $Ax = (1, -1, 0)$ do not have a solution, since $x = C(1, -1, 0) = (1/2, -1/2)$ does not satisfy $Ax = (1, -1, 0)$.

- The under-determined linear equations $A^T y = (1,2)$ has (different) solutions

$$B^T(1,2) = (1/3, 2/3, 38/9), \qquad C^T(1,2) = (0, 1/2, -1).$$

(Recall that B^T and C^T are both right inverses of A^T.) We can find a solution of $A^T y = b$ for any vector b.

Left and right inverse of matrix product. Suppose A and D are compatible for the matrix product AD (*i.e.*, the number of columns in A is equal to the number of rows in D.) If A has a right inverse B and D has a right inverse E, then EB is a right inverse of AD. This follows from

$$(AD)(EB) = A(DE)B = A(IB) = AB = I.$$

If A has a left inverse C and D has a left inverse F, then FC is a left inverse of AD. This follows from

$$(FC)(AD) = F(CA)D = FD = I.$$

11.2 Inverse

If a matrix is left- *and* right-invertible, then the left and right inverses are unique and equal. To see this, suppose that $AX = I$ and $YA = I$, *i.e.*, X is any right inverse and Y is any left inverse of A. Then we have

$$X = (YA)X = Y(AX) = Y,$$

i.e., any left inverse of A is equal to any right inverse of A. This implies that the left inverse is unique: If we have $A\tilde{X} = I$, then the argument above tells us that $\tilde{X} = Y$, so we have $\tilde{X} = X$, *i.e.*, there is only one right inverse of A. A similar argument shows that Y (which is the same as X) is the only left inverse of A.

When a matrix A has both a left inverse Y and a right inverse X, we call the matrix $X = Y$ simply the *inverse* of A, and denote it as A^{-1}. We say that A is *invertible* or *nonsingular*. A square matrix that is not invertible is called *singular*.

Dimensions of invertible matrices. Invertible matrices must be square, since tall matrices are not right-invertible, while wide matrices are not left-invertible. A matrix A and its inverse (if it exists) satisfy

$$AA^{-1} = A^{-1}A = I.$$

If A has inverse A^{-1}, then the inverse of A^{-1} is A; in other words, we have $(A^{-1})^{-1} = A$. For this reason we say that A and A^{-1} are inverses (of each other).