

LINEAR EQUATIONS

Systems of Linear Equations

A central task for linear algebra is to solve systems of linear equations. An example is to find values for unknowns that satisfy a set of linear equations:

$$\begin{array}{ll} x_1 + x_2 + x_3 = 5 & 1(1) + 1(2) + 1(2) = 5 \\ 2x_1 - 1x_2 + 6x_3 = 12 & 2(1) - 1(2) + 6(2) = 12 \\ x_1 + 3x_2 + 5x_3 = 17 & 1(1) + 3(2) + 5(2) = 17 \end{array}$$

The set of values, $x_1 = 1$; $x_2 = 2$; and $x_3 = 2$ solve each equation.

If this were the only set of three numbers that could simultaneously solve all three equations, then this system is said to have a *unique* solution.

In other systems of equations, there may be infinite solution sets or no solution.

A *consistent* system of equations has a solution.

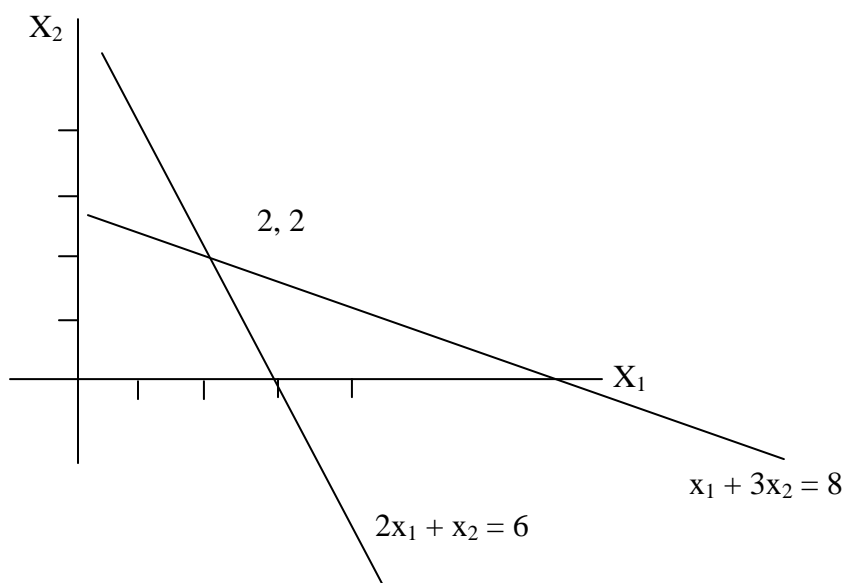
If only one solution is possible, it is *unique*

If an infinite set of solutions is possible, it is *dependent*

An *inconsistent* system of equations has no solution.

We learned how to solve equations graphically by identifying the point of intersection of two or more lines – the point of intersection provides the solution to the system of equations.

Consider: $2x_1 + x_2 = 6$ and $x_1 + 3x_2 = 8$



Consider: $x_1 + x_2 = 6$ and $2x_1 + 2x_2 = 12$

As you can imagine these form the same line. There are an infinite number of solutions for this system of equations. The equations are said to be *dependent* or *consistent*.

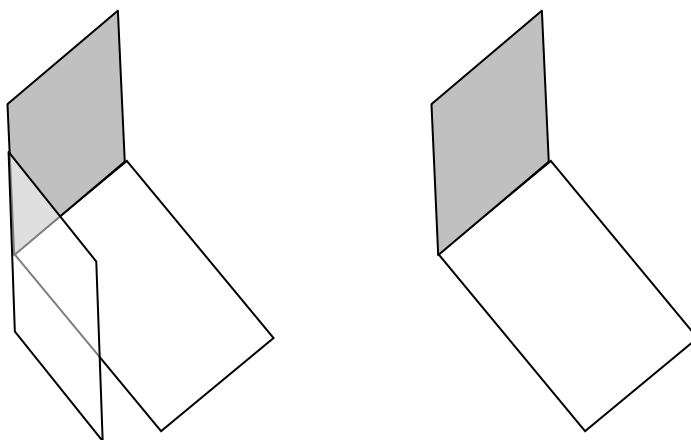
Consider: $2x_1 + x_2 = 6$

As you can also imagine, there is an infinite number of solutions to this single equation. The coordinates of any point on the line provide a solution to this equation.

Consider a system of three equations with three unknowns. Each equation can be represented by a plane in three-dimensions. To be consistent with a unique solution, the three planes must intersect at a single point in the system – typically creating a “corner.”

Just as in two-dimensional graphs, when two lines or two planes are parallel, the system is inconsistent where no solution is possible.

Consider the possibility that we have three equations with three unknowns – creating a system of three planes in space – where the three planes intersect at a point (or corner of a box). We remove one of the equations, leaving two planes that intersect on one side of each plane – the joint of the box. Now an infinite set of solutions is possible (but no unique solution).



→ Whenever we have more unknowns than equations, the system never has a unique solution. The system must be either dependent or inconsistent.

When there are more equations than unknowns, the system could be consistent and either provide unique or infinite solutions. It depends on the nature of the equations in the system.

Dependent Systems

Recall our two equations that result in the same line:

$$x_1 + x_2 = 6 \text{ and } 2x_1 + 2x_2 = 12$$

Through some algebra, we see that the second equation was derived by multiplying each term in the first equation by 2. This second equation is then *redundant*. Redundancy is usually not so easily seen.

Consider a separate set of equations from our first example.

$$x_1 + x_2 + x_3 = 5$$

$$2x_1 - 1x_2 + 6x_3 = 12$$

We could create redundancy by multiplying either equation by a non-zero scalar. Another way to create redundancy is to add the two equations together:

$$3x_1 + 7x_3 = 17$$

We could have multiplied the first two equations by a constant and then added them. In either case, the new equation is redundant but difficult to identify.

Inconsistent Equation Systems

An inconsistent system of equations can easily be created by multiplying the terms on one side of an equation by a number that is not equal to the multiple used on the other side of the equation.

Consider $x_1 + x_2 + x_3 = 5$

$$x_1 + x_2 + x_3 = 10, \text{ where we multiplied the left side by 1 and the right side by 2.}$$

Inconsistent systems share this characteristic, where one side has been disproportionately changed from the other side.

When a system of equations has a unique solution, no equation can be produced from any of the others through these kinds of operations – no equation in the system is redundant.

Matrix Characteristics

The Determinant

The determinant of matrix \mathbf{A} is denoted by $|\mathbf{A}|$ and is a unique scalar value associated with a square matrix. To give you an idea of its importance to statistics, the determinant of a variance-covariance matrix provides information about the *generalized* variance of several variables and thus gives us a useful tool for computing several multivariate statistics.

The determinant is difficult to compute once you move beyond a 3×3 matrix and those methods will not be discussed here. Computers can do this easily. For a 2×2 and 3×3 matrix, the process can be easily illustrated.

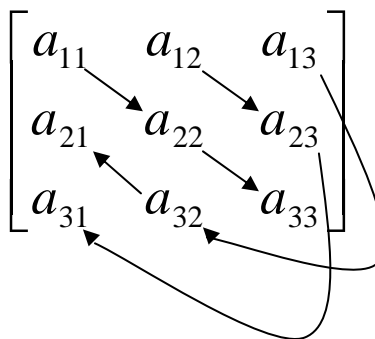
For a 2×2 matrix, the determinant can be found by:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

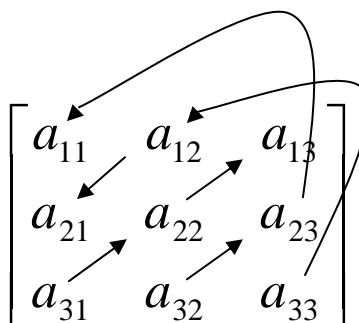
For a 3×3 matrix, the determinant can be found by:

$$|\mathbf{A}| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} \\ - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{12}a_{21}.$$

To obtain positive values:



To obtain negative values:



Consider an earlier example of two equations with a single solution:

$$2x_1 + x_2 = 6$$

$$x_1 + 3x_2 = 8$$

We can write this in matrix notation, including a vector of unknowns, a matrix of coefficients, and a vector of results (right hand of equations).

$$\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \quad \mathbf{A} \underline{x} = \underline{y}.$$

In this system, \mathbf{A} is a coefficient matrix, and would be used to compute the determinant.

We can use the determinant to evaluate a system of equations for unique solutions. If the determinant is zero, a unique solution does not exist.

Group Exercise:

What is the determinant of $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$? $|\mathbf{A}| =$

This system of equations has a unique solution.

Consider the equations: $x_1 + x_2 = 5$ and $2x_1 + 2x_2 = 5$

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \quad \mathbf{A} \underline{x} = \underline{y}.$$

What is the determinant of $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$? $|\mathbf{A}| =$

This system of equations does not have a unique solution.

When the determinant is equal to zero, we say that the system does not have a unique solution, the matrix is *singular*. $|\mathbf{A}| = 0$ if \mathbf{A} is singular.

When the determinant is not equal to zero, the matrix is *nonsingular* and there exists a unique solution.

Using SPSS Matrix command language to compute the determinant of matrix r:

COMPUTE determ = det(r).

The Inverse

Matrix algebra provides a succinct procedure to solve a system of linear equations. First, consider a simple algebraic problem:

$$5x = 25.$$

This is a problem with one unknown. Typically, we would divide each side by 5 to solve for x . An analogous method is to multiple each side by the reciprocal of 5, or $1/5$.

$$5^{-1} (5x) = 5^{-1} (25) \quad \rightarrow \quad x = 5$$

This is the process we use in matrix algebra. In order to complete this task with matrices, we must find the inverse of a matrix.

If we use our earlier matrix, $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, we can use the following steps to find the inverse of \mathbf{A} .

1. Computing the Inverse

- a. Find the determinant of the coefficient matrix.

$$|\mathbf{A}| = (2)(3) - (1)(1) = 5.$$

- b. Create a new matrix where we exchange the diagonal elements and make the off diagonals negative:

$$\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$$

- c. Multiply the converted matrix by the reciprocal of the determinant:

$$\mathbf{A}^{-1} = 1/5 \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{bmatrix}$$

2. Properties of the Inverse

- a. If the determinant of a matrix is zero, where the matrix is singular, the inverse does not exist.
- b. The inverse only exists for square matrices.
- c. If the inverse exists for a matrix, there is only one inverse.
- d. If a matrix is pre- or post-multiplied by its inverse, the product is the identity matrix: $\mathbf{A}^{-1} \mathbf{A} = \mathbf{A} \mathbf{A}^{-1} = \mathbf{I}$.
- e. The inverse provides us with information to solve a system of equations.

To complete our example, we now have the following:

$$\begin{aligned} \text{The equations:} \quad & 2x_1 + x_2 = 6 \\ & x_1 + 3x_2 = 8 \end{aligned}$$

$$\text{Matrix notation:} \quad \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \quad \mathbf{A} \underline{x} = \underline{y}.$$

$$\text{The determinant:} \quad |\mathbf{A}| = (2)(3) - (1)(1) = 5.$$

$$\text{The inverse:} \quad \mathbf{A}^{-1} = \begin{bmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{bmatrix}$$

$$\text{The solution:} \quad \mathbf{A} \underline{x} = \underline{y} \quad \rightarrow \quad \underline{x} = \mathbf{A}^{-1} \underline{y}$$

$$\begin{aligned} \underline{x} &= \begin{bmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 3/5(6) + -1/5(8) \\ -1/5(6) + 2/5(8) \end{bmatrix} = \begin{bmatrix} 18/5 - 8/5 \\ -6/5 + 16/5 \end{bmatrix} \\ &= \begin{bmatrix} 10/5 \\ 10/5 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

Inverse of a Diagonal Matrix

Find the inverse of the diagonal matrix \mathbf{S} .

$$\mathbf{S} = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \quad |\mathbf{S}| = (3)(5) - (0)(0) = 15$$

$$\mathbf{S}^{-1} = 1/15 \begin{bmatrix} 5 & -0 \\ -0 & 3 \end{bmatrix} = \begin{bmatrix} 5/15 & 0/15 \\ 0/15 & 3/15 \end{bmatrix} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/5 \end{bmatrix}$$

There is a general rule here: the inverse of a diagonal matrix is a diagonal matrix where the elements of the diagonal are the inverse of the elements of the original matrix.

Some general properties of the inverse:

$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ a matrix multiplied by its inverse is the identity matrix

$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ the inverse of the inverse is the original matrix

$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$ the inverse of a transpose is the transpose of the inverse

$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ the inverse of a product is the product of the inverses in reverse order

The inverse of a nonsingular matrix is unique.

If \mathbf{A} is nonsingular, then so is \mathbf{A}' .

Using SPSS matrix command language to compute the inverse of \mathbf{x} :

COMPUTE invx = inv(x).

An Application to Regression

Consider our earlier data set where we have scores on SAT, GPA, Self-Esteem and IQ for six individuals. We expect to predict SAT scores with great precision employing GPA, Self-Esteem, and IQ.

ID	SAT	GPA	Self-Esteem	IQ
1	560	3.0	11	112
2	780	3.9	10	143
3	620	2.9	19	124
4	600	2.7	7	129
5	720	3.7	18	130
6	380	2.4	13	82

One person's response to the explanatory variables can be stated in terms of an equation:

$$Y = b_1X_1 + b_2X_2 + b_3X_3 + E$$

The typical regression would try to estimate a solution to the set of equations where the b coefficients are computed in such a way to minimize the squared error (ordinary least squares). Now consider the possibility that we have 100 subjects. We would have 100 equations:

$$\begin{aligned} Y_1 &= b_{11}X_1 + b_{12}X_2 + b_{13}X_3 + E_1 \\ Y_2 &= b_{21}X_1 + b_{22}X_2 + b_{23}X_3 + E_2 \\ Y_3 &= b_{31}X_1 + b_{32}X_2 + b_{33}X_3 + E_3 \\ &\vdots \\ Y_{100} &= b_{100,1}X_1 + b_{100,2}X_2 + b_{100,3}X_3 + E_{100} \end{aligned}$$

We could denote this system of equations through matrix notation:

$$\underline{y} = \mathbf{X} \underline{b} + \underline{e}$$

with the dimensions: \underline{y} (100×1); \mathbf{X} (100×3); \underline{b} (3×1); \underline{e} (100×1);
so that $\mathbf{X} \underline{b}$ is conformable

Now we need to solve for \underline{b} .

Should we pre-multiply both sides of the equation by \mathbf{X}^{-1} ?

This will not work; because \mathbf{X} is not square, the inverse is undefined.

We need to address this by pre-multiplying both sides of the equation by the transpose of \mathbf{X} .

$$\mathbf{X}' \underline{y} = \mathbf{X}'(\mathbf{X} \underline{b} + \underline{e}) = \mathbf{X}' \mathbf{X} \underline{b} + \mathbf{X}' \underline{e} = \mathbf{X}' \mathbf{X} \underline{b}$$

This is the case since \underline{e} is a vector defined to be orthogonal to our three \underline{x} vectors.

Note that the product matrix $\mathbf{X}'\mathbf{X}$ is a columns product matrix; a square (3×3) matrix.

We now have: $\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\mathbf{b}$

We can now premultiply each side by the inverse of the columns product matrix:

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\mathbf{b}$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{I}\mathbf{b}$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{b}$$

The solution for regression weights is possible as long as $\mathbf{X}'\mathbf{X}$ is nonsingular and the inverse exists.

This solution assumes that the variables employed are all standardized (mean = 0, sd = 1). The resulting weights are standardized weights.

To solve for unstandardized regression weights, we can employ the typical equation, which includes the intercept to account for differences in variable means.

$$Y_1 = b_{10} + b_{11}X_1 + b_{12}X_2 + b_{13}X_3 + E_1$$

which can be summarized in matrix notation by

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{100} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{100,1} & x_{100,2} & x_{100,3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{100} \end{bmatrix}$$

Note that the dimensions are as expected: $(100 \times 1) = (100 \times 4)(4 \times 1) + (100 \times 1)$.

The solution provided earlier applies to this system similarly:

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{b}$$

The Inverse and Regression

In regression analysis, the matrix that must be inverted is $\mathbf{X}'\mathbf{X}$.

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$|\mathbf{X}'\mathbf{X}| = n \sum X_i^2 - (\sum X_i)(\sum X_i)$$

$$= n \sum X_i^2 - (\sum X_i)^2$$

$$= n \sum X_i^2 - (n\bar{X})^2$$

$$= n \sum (X_i - \bar{X})^2$$