Basics of Linear Algebra

Alex Houtz[*]

August 10, 2023

Linear algebra is the study of linear systems of equations. In economics, we often have systems of equations that we want to solve. Naturally, then, linear algebra is a key tool of choice for us. Like yesterday, we will begin with the basics.

1 Basics Concepts

A **vector** $x \in \mathbb{R}^n$ looks as follows:

$$x = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}$$

Each $x_i \in \mathbb{R}$. When a vector is horizontal, as above, we call it a row vector. When the vector is vertical, we call it a column vector. We can add two vectors or multiply a vector by a scalar as follows:

$$x + y = \begin{bmatrix} x_1 + y_1 & x_2 + y_2 & \dots & x_n + y_n \end{bmatrix}$$
$$\alpha x = \begin{bmatrix} \alpha x_1 & \alpha x_2 & \dots & \alpha x_n \end{bmatrix}$$

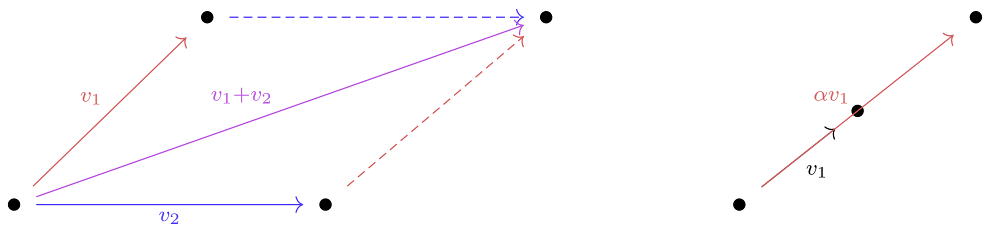
We can represent these two concepts graphically (see figure on the next page).

The **inner product** of two vectors is defined as:

$$x \cdot y = x_1 y_1 + x_2 y_2 + \dots x_n y_n$$

We also call the inner product the dot product.

[*]Math Camp Instructor | University of Notre Dame



It is often useful to define a distance measure, or **norm**, in a vector space. A norm must have absolute homogeneity, satisfy the triangle inequality, and be positive definite. In other words, for p to be a norm in vector space X:

1. $p(sx) = |s|p(x)$ for all $x \in X$ and all $s \in \mathbb{R}$
2. $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$
3. If $p(x) = 0$, then $x = 0$ for all $x \in X$

Our vector space is $\mathbb{R}^n$. The norm on $\mathbb{R}^n$ is the Euclidean norm:

$$\|x\| = \sqrt{x \cdot x}$$

We can use the norm to find the angle, θ, formed between two vectors:

$$x \cdot y = \|x\|\|y\| \text{cos}\theta$$

Note that this formulation necessarily implies that if x and y are orthogonal, $\theta = \pi/2$ radians.

A **matrix**, A, is an array of row and column vectors:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

The **row space** of a matrix is the vector space in which the row vectors of the matrix live. In A, the row space is $\mathbb{R}^m$. The **column space** is the vector space in which the column vectors of the matrix live. In A, this is $\mathbb{R}^n$. We say a matrix is **square** if the row space has the same dimension as the column space.

Like vectors, we can add two matrices of the same dimensions:

$$A \pm B = a_{ij} + b_{ij}$$

Or multiply by a scalar:

$$\alpha A = [\alpha a_{ij}]$$

We can also multiply two vectors together if the column space of the first matrix matches the row space of the second matrix:

$$A_{m \times n} B_{n \times k} = AB_{m \times k} = \begin{bmatrix} a_{1j}b_{i1} & a_{1j}b_{i2} & \dots & a_{1j}b_{ik} \\ a_{2j}b_{i1} & a_{2j}b_{i2} & \dots & a_{2j}b_{ik} \\ a_{mj}b_{i1} & a_{mj}b_{i2} & \dots & a_{mj}b_{ik} \end{bmatrix}$$

We can deduce the following laws of matrix algebra from these three definitions:

- Associative Law: $(A + B) + C = A + (B + C)$ and $(AB)C = A(BC)$
- Commutative Law: $A + B = B + A$. Note that this law does not hold for multiplication.
- Distributive Law: $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$

Sometimes we want to swap the row and column space of the matrix. This is called the **transpose** of the matrix, denoted by A^T or A' Thus, a_{ij} becomes a_{ji}. Here are some rules for transposing a matrix:

- $(A \pm B)' = A' \pm B'$
- $(A')' = A$
- $(\alpha A)' = \alpha A'$
- $(AB)' = B'A'$

This last rule will be useful throughout the year more often than you might expect.

You'll use a variety of matrix types throughout the year and in your own research. Here are a few of the most common types of matrices:

- Identity, I
- Diagonal
- Inverse, $AA^{-1} = I$
- Upper/Lower triangular
- Symmetric
- Toeplitz

Inverses also have useful properties:

- $(A')^{-1} = (A^{-1})'$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $|A|^{-1} = |A^{-1}|$
- $(cA)^{-1} = \frac{1}{c}A^{-1}$

1.1 Derivatives

Finally, we can do differential calculus on matrices and vectors. I break this into five cases.

Vector by Scalar

We can differentiate a vector by a scalar. Let y be a column vector and x be a scalar. Then:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \dots & \frac{\partial y_m}{\partial x} \end{bmatrix}'$$

You may have seen this derivative referred to as the **tangent vector** of y.

Matrix by Scalar

Let Y be a matrix and x be a scalar. Then:

$$\frac{\partial Y}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \dots & \frac{\partial y_{1n}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \dots & \frac{\partial y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \frac{\partial y_{m2}}{\partial x} & \dots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix}$$

Scalar by Vector

Let y be a vector and x be a scalar. Then:

$$\frac{\partial x}{\partial y} = \begin{bmatrix} \frac{\partial x}{\partial y_1} & \frac{\partial x}{\partial y_2} & \dots & \frac{\partial x}{\partial y_m} \end{bmatrix}'$$

You may have seen this derivative referred to as the **gradient**, particularly when applied to a function and denoted by ∇f.

Vector by Vector

Let both x and y be vectors. Then:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

This derivative is called the **Jacobian** and sometimes denoted by $D_x y$.

Scalar by Matrix

Let Y be a matrix and x a scalar. Then:

$$\frac{\partial x}{\partial y} = \begin{bmatrix} \frac{\partial x}{\partial y_{11}} & \frac{\partial x}{\partial y_{12}} & \cdots & \frac{\partial x}{\partial y_{1n}} \\ \frac{\partial x}{\partial y_{21}} & \frac{\partial x}{\partial y_{22}} & \cdots & \frac{\partial x}{\partial y_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x}{\partial y_{m1}} & \frac{\partial x}{\partial y_{m2}} & \cdots & \frac{\partial x}{\partial y_{mn}} \end{bmatrix}$$

Note that we can take the derivative of a vector by a matrix or a matrix by a matrix. These rely on tensors though and are beyond the scope of our coverage here.

1.2 Linear Independence and Determinant

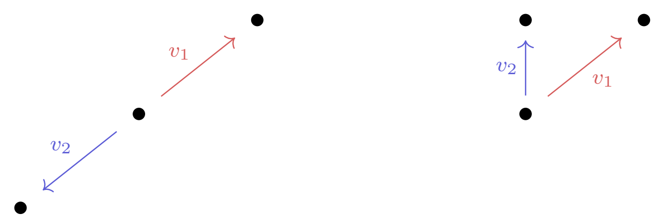
Suppose that we have a collection of vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$. These vectors are **linearly independent** if $\alpha_i = 0$ is the unique solution to:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

Otherwise, these vectors are linearly dependent. Note that we can write this equation in matrix and vector form:

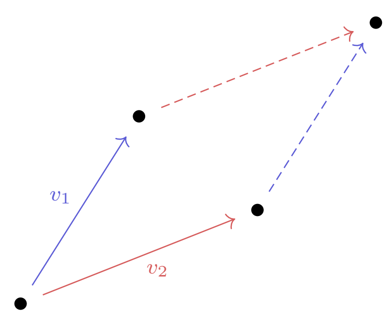
$$\boldsymbol{\alpha v} = \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1k} \\ v_{21} & v_{22} & \ldots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \ldots & v_{nk} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$$

The **rank** of the matrix is the number of linearly independent columns (or, equivalently, rows). The set of linearly independent vectors, V, is said to **span** a set $X \subset \mathbb{R}^n$ if any vector $x \in X$ can be expressed as a linear combination of vectors in V. A span can have more than just linearly independent vectors in X, but it must have all linearly independent vectors in X. V forms a basis for $\mathbb{R}^n$ if every vector in V is linearly independent and spans $\mathbb{R}^n$. Graphically:



The left picture shows two linearly dependent vectors. v_2 is simply $(-1)v_1$. Thus, in the left picture, v_1 and v_2 span the line they are both on. But in the right picture, v_1 and v_2 are linearly independent. Together, they can span all of $\mathbb{R}^n$, forming a basis for $\mathbb{R}^n$.

The **determinant** of a matrix summarizes lots of information about a matrix in one number. The mathematical definition is somewhat technical and does not provide a lot of intuition, so we'll focus on the geometric meaning of the determinant in $\mathbb{R}^2$. Take the column vectors of a 2×2 matrix and add them together to generate the following picture:



The absolute value of the determinant is the area inside the parallelogram formed by the column vectors. As the column vectors get closer to being linearly dependent, the area inside the parallelogram goes to zero. In econometrics, as the determinant goes to zero, we cannot retrieve regression coefficients. We say that we have a multicolinearity problem.

How do we actually calculate determinants? Take a 3×3 matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Using a Laplace expansion, the determinant of A, denoted as $|A|$, is given by:

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} m_{ij}$$

where m_{ij} is the determinant of the submatrix given by removing row i and column j of A. For us, expanding along the first row:

$$|A| = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2} a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3} a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

This method generalizes to any $n \times n$ matrix, though as n becomes large, using a method from Gaussian elimination proves faster.

1.3 Solving Linear Systems

We now have all the tools we need to solve linear problems. Suppose we have a linear system of the form:

$$Ax = b$$

We want to find x. Let A be an $n \times n$ matrix. Then all of the following statements are equivalent:

1. A is invertible iff $|A| \neq 0$
2. For any $b \in \mathbb{R}^n$, $Ax = b$ has a unique solution.
3. A is full rank
4. All column/row vectors of A are linearly independent
5. Column vectors form a basis for $\mathbb{R}^n$

We know, then, that if our matrix is of full rank, or if the determinant is not zero, or if all the column vectors of A are independent, we can invert A to find x.

When solving systems by hand, a useful rule to know is **Cramer's rule**. Cramer's rule states that we can find x_i using;

$$x_i = \frac{|A_i|}{|A|}$$

where A_i is formed by replacing the i^{th} column of A with vector b. Suppose we want to find x_2 from the 3×3 example above. Using Cramer's rule:

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_2 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{|A|}$$

1.4 Quadratic Forms

A quadratic form is defined as:

$$Q(x) = x'_{1\times n} A_{n\times n} x_{n\times 1}$$

In the case where $n = 2$, the quadratic form is:

$$Q(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Notice that A must be symmetric. If we multiplied this out, we would get:

$$Q(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

which looks like the standard quadratics we've all dealt with many times. When we optimize a quadratic normally, we want to find the solution to the first-order condition and then find whether that solution is a maximum or minimum. To do this with systems, we need leading principle minors to determine whether a multivariate function is concave or convex. To find leading principle minors, first complete the square on the quadratic:

$$Q(x) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}x_2}\right)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}}x_2^2$$

Define $M_1 \equiv a_{11}$ and let $x_1 + \frac{a_{12}}{a_{11}x_2} \equiv y_1$. Let $M_2 \equiv a_{11}a_{22} - a_{12}^2$ and $y_2 \equiv x_2$. This yields:

$$Q = M_1 y_1^2 + \frac{M_2}{M_1}y_2^2$$

For $Q > 0$, regardless of the value of y_1 or y_2, M_1 and M_2 must be greater than 0. For $Q < 0$, regardless of the value of y_1 or y_2, $M_1 < 0$ and $M_2 > 0$. But M_1 and M_2 are entirely composed of elements of A, so we can calculate them. In general, quadratics can be expressed as:

$$Q = M_1 y_1^2 + \frac{M_2}{M_1}y_2^2 + \cdots + \frac{M_n}{M_{n-1}}y_n^2$$

where:

$$\begin{aligned}
M_1 &= |a_{11}| \\
M_2 &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\
M_3 &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&\vdots
\end{aligned}$$

We call each M_i a **leading principle minor** (LPM). LPMs tell us that:

1. A is positive definite iff all LPMs are positive – A is positive definite iff Q $>$ 0 for all $x \neq 0$.

2. A is negative definite iff all LPMs alternate in sign, starting with $M_1 < 0$ – A is negative definite iff $Q < 0$ for all $x \neq 0$.

There are three other types of definiteness:

1. A is positive semi-definite iff all principal minors are non-negative – A is positive semi-definite iff $Q \geq 0$ for all $x \neq 0$.

2. A is negative semi-definite iff all principal minors alternate in sign, starting with the first-order principal minors < 0 – A is negative semi-definite iff $Q \leq 0$ for all $x \neq 0$.

3. If a matrix does not match any of the above definitions, it is indefinite.

A standard **principal minor** is found by eliminating column $i =$ row j and calculating the determinant of the leftover terms. In a 3×3 matrix, the first-order principal minors require us to eliminate two rows and columns. The second-order principal minors require us to eliminate one row and column. The third-order principal minor is equivalent to the third LPM. Denote the principal minors as Δ_k, where k is the order of the principal minor.

Example 1. Suppose matrix A looks like:

$$\begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 2 \\ 6 & 1 & 6 \end{bmatrix}$$

Then the leading principal minors are:

$$M_1 = 1, \quad M_2 = \begin{vmatrix} 1 & 4 \\ 4 & 2 \end{vmatrix} = -14, \quad M_3 = \begin{vmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{vmatrix} = -109$$

$$\Delta_1 = 1, 2, 6 \quad \Delta_2 = -14$$

The leading principal minors don't match either positive nor negative definiteness. The principal minors don't match either positive nor negative semi-definiteness. Therefore, A is indefinite.

1.5 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are extremely important in economic models, governing whether the model has an equilibrium and is stable. Why? Consider a first order difference equation, which can represent most macro models, and which looks like:

$$x_{t+1} = Ax_t + b$$

where x is a vector, A is a matrix, and b is a vector. There's something called a steady-state (where $x_{t+1} = x_t$). We often want to study what happens when the economy is shocked out of steady-state. To do so, we often need to diagonalize A. We do this by conducting a spectral decomposition (if A is symmetric), or eigenvalue decomposition. If the eigenvalues are too large, the system will explode. If all are too small, we could have many paths back to steady-state. If we have the right mix, we will get one "equilibrium" path back to steady-state.

First, let's focus on the basics of eigenvalues and eigenvectors. Consider matrix A linearly transforming vector v. Then the **eigenvalue** of matrix A is the λ such that:

$$Av = \lambda v$$

where v is the **eigenvector**. Rearranging this equation yields:

$$(A - \lambda I)v = 0$$

This equation has a non-trivial solution iff $|A - \lambda I|$ is zero. Why? Because the eigenequation demands that Av and λv are colinear – not linearly independent. We call $|A - \lambda I|$ the **characteristic polynomial** of A. Let's do an example.

Example 2. Suppose that A takes the following form:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Then the characteristic polynomial is:

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} \\ &= 3 - 4\lambda + \lambda^2 \end{aligned}$$

Set this equal to zero and solve the quadratic:

$$3 - 4\lambda + \lambda^2 = (\lambda - 3)(\lambda - 1) = 0$$

So $\lambda = 1, 3$ are the two eigenvalues of A. To find the eigenvectors, plug each value of λ into the eigenequation and solve for v: For $\lambda = 1$:

$$\begin{aligned} 0 &= (A - \lambda I)v_{\lambda=1} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &\Rightarrow 1v_1 + 1v_2 \\ v_1 &= -v_2 \end{aligned}$$

Any values that satisfy this equation will be an eigenvector for $\lambda = 1$. For simplicity, I'll just choose:

$$v_{\lambda=1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

We can similarly find that:

$$v_{\lambda=3} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

After finding all the eigenvalues and linearly independent eigenvectors, we can do an **eigendecomposition** on $A_{n \times n}$:

$$A = V \Lambda V^{-1}$$

where V is an $n \times n$ matrix with column i being eigenvector v_i and Λ is a diagonal matrix whose diagonal elements are the corresponding eigenvalues ($\Lambda_{ii} = \lambda_i$). Usually we orthonormalize the eigenvectors, but we don't have to. Note that if each eigenvector has a unique eigenvalue, all the eigenvectors will be orthogonal. This condition is always satisfied if A is a symmetric matrix.

Continuing the example above, we can eigendecompose A:

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

Multiplying out the matrices confirms the decomposition. Now, let's substitute the decomposition in for A into a system:

$$\begin{aligned} Ax &= b \\ V\Lambda V^{-1}x &= b \\ \Lambda V^{-1}x &= V^{-1}b \end{aligned}$$

Define $\tilde{x} \equiv V^{-1}x$ and $\tilde{b} \equiv V^{-1}b$. Then:

$$\Lambda\tilde{x} = \tilde{b}$$

Note that Λ is a diagonal matrix, which makes this equation very easy to solve:

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}$$

So $\tilde{x}_1 = \tilde{b}_1$ and $\tilde{x}_2 = \tilde{b}_2/3$. We call systems where $\lambda_i\tilde{x}_i = \tilde{b}_i$ **uncoupled** systems. Of course, we don't want to find $\tilde{x}$, we want to find x. Premultiply by V to get back x.

Let's now return to our first-order difference equation:

$$x_{t+1} = Ax_t + b$$

First, let's find the steady-state:

$$x_{ss} = (I - A)^{-1}b$$

Then any deviation from the steady-state will be:

$$x_{t+1} = x_{ss} + A(x_t - x_{ss})$$

Substitute in the eigendecomposition for A:

$$x_{t+1} = x_{ss} + V\Lambda V^{-1}(x_t - x_{ss})$$

Notice that this system is now uncoupled. Define $z_t = V^{-1}(x_t - x_{ss})$. Then we can rewrite the system as:

$$z_{t+1} = \Lambda z_t$$

Recursively substitute in for z_t to get:

$$z_{t+1} = \Lambda^t z_0$$

Or, written in the original formulation:

$$x_{t+1} = x_{ss} + V\Lambda^t V^{-1}(x_0 - x_{ss})$$

where x_0 is the shocked value of x. Now, as $t \to \infty$, Λ^t either explodes or goes to zero. The stability of the system thus depends on the value of the eigenvalues.

Lastly, let's look at the quadratic form. Substitute in the eigendecomposition:

$$x'Ax = x'V\Lambda V^{-1}x$$

Recall that in the quadratic form, A is symmetric. Therefore, $V^{-1} = V'$.[1] Define $y = V'x$. Then the quadratic form will become:

$$\begin{aligned} x'Ax &= y'\Lambda y \\ &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \end{aligned}$$

This new form leads to four results. Symmetric matrix A is:[2]

1. Positive definite iff all eigenvalues are strictly positive.

[1] The eigenvectors are all orthogonal due to A being symmetric. By definition, a square matrix, B, is orthogonal if $BB' = B'B = I$. Premultiply by B^{-1} to get:

$$B^{-1}BB' = B^{-1}I \Rightarrow B' = B^{-1}$$

[2] It should be noted that there are a great many properties concerning positive and negative (semi)definiteness. Simon and Blume is a good resource for learning more.

2. Negative definite iff all eigenvalues are strictly negative.
3. Positive semi-definite iff all eigenvalues are non-negative.
4. Negative semi-definite iff all eigenvalues are non-positive.

In addition, if A is a symmetric matrix, the following three statements are equivalent:

1. A is positive definite
2. $\exists$ a non-singular (invertible) matrix B such that $A = B'B$
3. $\exists$ a non-singular matrix Q such that $Q'AQ = I$