Eigenvalues and Dynamics

In our last meeting we used the notions of first-order critical value and the second derivative test to identify extrema.

In today’s class we will use matrix algebra to examine multivariate problems.

We will also discuss another important application of matrix algebra: the solution of Markov processes.

Multivariate Problems

Many problems in the social sciences involve the extension of some familiar results from calculus to a multivariate setting.

Matrix algebra plays an important role in these cases, because the second order condition for minimization or maximization problems with several variables requires that one check the signs of the discriminant of certain matrices of second derivatives.

So, let’s go back to our discussion of matrix algebra. A few weeks ago, we learned that:

The rank of a square matrix determines whether the matrix has an inverse.

A square matrix of order $n \times n$ is said to be of full rank if its rank is $n$.

A square matrix of full rank has an inverse; such matrices are said to be nonsingular.

Square matrices with less than full rank are said to be singular, and they are not invertible.

We also learned how to use Gaussian elimination to compute the inverse of a square matrix, if it exists.

Suppose we are given the following matrix:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
we can use the identity matrix, $I_2$ to obtain the matrix

$$[A|I] = \begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix}$$

If $a$ and $c$ are both 0, $A$ will clearly be singular. Let us assume, then, that $a \neq 0$. We can now perform the following operations to obtain the row echelon form of this matrix.

$$\begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix} \rightarrow -\frac{c}{a}R_1 + R_2 \rightarrow \begin{bmatrix} a & b \ \frac{ad-bc}{a} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$ 

Note that when $a \neq 0$, $A$ is nonsingular (and therefore invertible) if and only if $ad-bc \neq 0$.

We can now perform the following operations to obtain the reduced row echelon form of this matrix.

$$\begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix} \rightarrow \frac{1}{a}R_1 \rightarrow \begin{bmatrix} 1 & b/a \ \frac{ad-bc}{a} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{c}{a}.$$ 

$$\begin{bmatrix} 1 & b/a & 0 & 1 \\ 0 & 1 & -\frac{a}{ad-bc} \end{bmatrix} \rightarrow \frac{1}{a}R_2 + R_1 \rightarrow \begin{bmatrix} 1 & 0 \ \frac{ad-bc}{a} \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$ 

$$A^{-1} = -\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$ 

Note that if $ad-bc \neq 0$, then $a$ and $c$ cannot be both 0.

Therefore, we have prove the following theorem:

**Theorem 1** *The general $2 \times 2$ matrix given by $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is nonsingular (and therefore invertible) if and only if $ad-bc \neq 0$. Its inverse is $A^{-1} = -\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$"
Let us take a look at the simplest of $n \times n$ matrices: a $1 \times 1$ matrix. Such matrix is just a scalar ($a$).

Since the inverse of $a$, $\frac{1}{a}$ exists if and only if $a$ is nonzero, it is natural to define the determinant of such matrix to be just that scalar $a$:

$$\text{det}(a)=a$$

For a $2 \times 2$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

the Theorem that we just proved states that $A$ is nonsingular if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Therefore, we define the determinant of a $2 \times 2$ matrix $A$:

$$\text{det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$  

Notice that in this case, the determinant is the product of the two diagonal entries minus the product of the two off-diagonal entries.

So, we can rewrite the same expression as:

$$\text{det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \text{det}(a_{22}) - a_{12} \text{det}(a_{21}).$$

The first term on the RHS of this last expression is the $(1,1)^{th}$ entry of $A$ times the determinant of the submatrix obtained by deleting from $A$ the row and column which contain that entry.

The second term is the $(1,2)^{th}$ entry times the determinant of the submatrix obtained by deleting from $A$ the row and column which contain that entry.

The terms alternate in sign; the term containing $a_{11}$ receives a plus sign and the term containing $a_{12}$ receives a minus sign.

The following definitions will simplify the task of defining the determinant of an $n \times n$ matrix.

Let $A$ be an $n \times n$ matrix. Let $A_{ij}$ be the $(n-1) \times (n-1)$ submatrix obtained by deleting row $i$ and column $j$ from $A$.

The determinant of $A_{ij}$ is called a **minor** of $A$, $M_{ij} \equiv \text{det} A_{ij}$.  

3
When the correct sign, \((-1)^{i+j}\), is applied, it becomes a **cofactor**, \(C_{ij} \equiv (-1)^{i+j}M_{ij}\).

Note that \(M_{ij} = C_{ij}\) if \((i + j)\) is even and \(M_{ij} = -C_{ij}\) if \((i + j)\) is odd.

We can now write the determinant of a \(3 \times 3\) matrix as

\[
\begin{vmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}
= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}
= a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.
\]

The \(j^{th}\) term on the RHS of the expression is \(a_{1j}\) times the determinant of the submatrix obtained by deleting row 1 and column \(j\) from \(A\).

The term is preceded by a plus sign if \(1 + j\) is even and by a minus sign if \(1 + j\) is odd.

**Definition 1** The determinant of an \(n \times n\) matrix \(A\) is given by

\[
\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}
= a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}
\]

In referring to the determinant of an \(n \times n\) matrix \(A\), one sometimes writes

\[
\begin{vmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
\]

for \(\det\)

\[
\begin{vmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
\]

and \(|A|\) for \(\det A\).

Note that our definition of the determinant of a matrix involves expanding along its first row. This operation, though, can be done using any other row or any other column as well.

In any case, computing the determinant of an \(n \times n\) is rather tedious (it involves \(n!\) terms, each a product of \(n\) entries).

It is unlikely that any of you guys will ever calculate any determinants over \(3 \times 3\) without a computer.

A \(3 \times 3\), however, might be computed on occasion; if so, the following “shortcut” will prove useful:
\[
\begin{vmatrix}
  a_{11} & a_{12} & a_{1n} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{22}a_{31} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{23}a_{32}.
\]

**Uses of the Determinant.**

Since the determinant tells us whether or not \(A^{-1}\) exists, we can also use it to determine whether or not \(Ax = b\) has a unique solution.

Recall that if \(A\) is an \(n \times n\) matrix, and \(x\) and \(b\) are both column vectors having \(n\) elements, the former consisting of unknowns and the latter of known constants, then a solution to the system of equations \(Ax = b\) can be obtained by premultiplying both sides of the equation by \(A^{-1}\), if it exists.

This is so because

\[A^{-1}Ax = A^{-1}b\]

is the same as

\[Ix = A^{-1}b\]

by virtue of the definition of the inverse \((A^{-1}A = I)\), and this in turn is the same as

\[x = A^{-1}b\]

by virtue of the definition of the identity matrix.

Consider now the calculation of the inverse matrix. For a \(2 \times 2\) matrix, \(AB = I\) implies that

\[
\begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
  b_{11} & b_{12} \\
  b_{21} & b_{22}
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 \\
  0 & 1
\end{bmatrix}
\]

or

\[
a_{11}b_{11} + a_{12}b_{21} = 1 \\
a_{11}b_{12} + a_{12}b_{22} = 0 \\
a_{21}b_{11} + a_{22}b_{21} = 0 \\
a_{21}b_{12} + a_{22}b_{22} = 1
\]

The solutions are

\[
\begin{bmatrix}
  b_{11} & b_{12} \\
  b_{21} & b_{22}
\end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}}
\begin{bmatrix}
  a_{22} & -a_{12} \\
  -a_{21} & a_{11}
\end{bmatrix}
\]

5
or

\[ A^{-1} = \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \]

Notice the presence of the reciprocal of \( |A| \) in \( A^{-1} \).

For any \( n \times n \) matrix \( A \), let \( C_{ij} \) denote the \((i, j)\)th cofactor of \( A \). The \( n \times n \) matrix whose \((i, j)\)th entry is \( C_{ji} \), the \((j, i)\)th cofactor of \( A \) (note the switch in indices), is called the **adjoint** of \( A \) and is written \( \text{adj} \ A \).

The adjoint matrix of \( A \) is the transpose of the matrix of cofactors of \( A \).

**Example 2** Consider the following matrix, \( A = \begin{bmatrix} 2 & 4 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix} \)

We can find \( A^{-1} \) using the determinant:

\[
C_{11} = + \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = 3, \quad C_{12} = - \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = 0, \quad C_{13} = + \begin{bmatrix} 0 & 3 \\ 1 & 0 \end{bmatrix} = -3, \\
C_{21} = - \begin{bmatrix} 4 & 5 \\ 0 & 1 \end{bmatrix} = -4, \quad C_{22} = + \begin{bmatrix} 2 & 5 \\ 1 & 1 \end{bmatrix} = -3, \quad C_{23} = - \begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix} = 4, \\
C_{31} = + \begin{bmatrix} 4 & 5 \\ 3 & 0 \end{bmatrix} = -15, \quad C_{32} = - \begin{bmatrix} 2 & 5 \\ 0 & 0 \end{bmatrix} = 0, \quad C_{33} = + \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} = 6,
\]

It is easy to see that \( |A| = -9 \).

We can also easily calculate \( \text{adj} \ A \).

\[
\text{adj} \ A = \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = \begin{bmatrix} 3 & -4 & -15 \\ 0 & -3 & 0 \\ -3 & 4 & 6 \end{bmatrix},
\]

So,

\[
A^{-1} = -\frac{1}{9} \begin{bmatrix} 3 & -4 & -15 \\ 0 & -3 & 0 \\ -3 & 4 & 6 \end{bmatrix}.
\]

Now we are ready to find the unique solution to a system of \( n \) linear equations in \( n \) unknowns, \( x_1, x_2, \ldots, x_n \).
We can write this system of equations in matrix form as $Ax = b$, where

$$\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Cramer’s rule. To find the component $x_i$ of the solution vector to this system of linear equations, replace the $i$th column of the matrix $A$ with the column vector $b$ to form a matrix $A_i$. Then $x_i$ is the determinant of $A_i$ divided by the determinant of $A$: $x_i = \frac{|A_i|}{|A|}$

Example 3 Consider the following system of equations,

$$\begin{bmatrix} 1 & 1 & 1 \\ 12 & 2 & -3 \\ 3 & 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ -4 \end{bmatrix}.$$

We can use Cramer’s rule to calculate $x_3$:

The determinant of the coefficient matrix $A$ is 35.

The determinant of $A_3 = \begin{bmatrix} 1 & 1 & 0 \\ 12 & 2 & 5 \\ 3 & 4 & -4 \end{bmatrix}$ is also 35.

Therefore, $x_3 = \frac{|A_3|}{|A|} = 1$.

Definite and semidefinite matrices.

Let $A$ be a symmetric square matrix. Then, if we postmultiply $A$ by some vector $x$ and premultiply it by the (transpose of the) same vector $x$, we have a quadratic form. For example,

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + (a_{21} + a_{12})x_1x_2 + a_{22}x_2^2.$$

Suppose that $A$ is the identity matrix.

In this case, it is not hard to see that whatever the values of $x_1$ and $x_2$, the quadratic form must be nonnegative.
In fact, if \( x_1 \) and \( x_2 \) are not both zero, \( x'Ax \) will be strictly positive.

The identity matrix is an example of a **positive definite matrix**.

**Definition 1** A square matrix \( A \) is

(a.) **positive definite** if \( x'Ax > 0 \) for all \( x \neq 0 \);

(b.) **negative definite** if \( x'Ax < 0 \) for all \( x \neq 0 \);

(c.) **positive semidefinite** if \( x'Ax \geq 0 \) for all \( x \);

(d.) **negative semidefinite** if \( x'Ax \leq 0 \) for all \( x \).

In some cases we do not want to require that \( x'Ax \) has a definite sign for all values of \( x \), but only for some restricted set of values.

We say \( A \) is **positive definite subject to the constraint** \( bx = 0 \) if \( x'Ax > 0 \) for all \( x \neq 0 \) such that \( bx = 0 \).

The other definitions extend to the constrained case in a natural manner.

**Eigenvalues and Eigenvectors**

Associated with a square matrix are numbers called **eigenvalues** and vectors called **eigenvectors**.

The eigenvalues of a given \( n \times n \) matrix are the \( n \) numbers which summarize the essential properties of that matrix.

Since these \( n \) numbers characterize the matrix under study, they are often called the “characteristic values” of the matrix.

They are commonly referred to by the German term “eigenvalues.”

Consider a square matrix \( A \). A useful set of results for analyzing such matrix arises from the solutions to the set of equations

\[
Ac = \lambda c.
\]

The pairs of solutions are the **characteristic vectors**, or eigenvectors, \( c \) and characteristic roots, or eigenvalues, \( \lambda \).

If \( c \) is any solution vector, \( kc \) is also for any value of \( k \).

To remove the indeterminacy, \( c \) is **normalized** so that the sum of squares \( c_1^2 + \cdots + c_n^2 \) equals one.
\[ c'c = 1. \]

The solution then consists of \( \lambda \) and the \( n - 1 \) unknown elements in \( c \).

**The Characteristic Equation.** Solving \( Ac = \lambda c \) can, in principle, proceed as follows.

First, we can rewrite the expression as

\[ Ac = \lambda Ic \]

which implies that

\[ (A - \lambda I)c = 0. \]

This is a homogeneous system. Recall that any homogeneous system has either a unique solution (i.e. the *trivial solution*), or many solutions.

The latter is true if the matrix \( (A - \lambda I) \) is singular or has a zero determinant.

Therefore, if \( \lambda \) is a solution,

\[ |A - \lambda I| = 0. \]

This polynomial in \( \lambda \) is the **characteristic equation** of \( A \).

**Example 2** Consider the following matrix, \[
\begin{bmatrix}
 5 & 1 \\
 2 & 4 \\
\end{bmatrix}.
\]

Then \( |A - \lambda I| = \begin{vmatrix} 5 - \lambda & 1 \\ 2 & 4 - \lambda \end{vmatrix} = (5 - \lambda)(4 - \lambda) - 2(1) = \lambda^2 - 9\lambda + 18. \)

The two solutions are \( \lambda = 6 \) and \( \lambda = 3 \).

In solving the characteristic equation, there is no guarantee that the characteristic roots will be real.

In the preceding example if the 2 in the lower right-hand corner of the matrix were -2 instead, the solution would be a pair of complex values.

The same problem can emerge in the general \( n \times n \) case.

Nonetheless, the characteristic roots of a symmetric matrix are real. This result is very convenient because a lot of applications involve the characteristic roots and vectors of symmetric matrices.
For an $n \times n$ matrix, the characteristic equation is an $n$th-order polynomial in $\lambda$. Its solutions may be $n$ distinct values, as in the preceding example, or may contain repeated values of $\lambda$, as in

$$\begin{vmatrix} 2 - \lambda & 0 \\ 0 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 \Rightarrow \lambda_1 = \lambda_2 = 2,$$

and may contain some zeroes as well, as in

$$\begin{vmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda = 0 \Rightarrow \lambda_1 = 5 \text{ and } \lambda_2 = 0.$$

**Eigenvectors.** With $\lambda$ in hand, the characteristic vectors are derived from the original problem,

$$Ac = \lambda c.$$

or

$$(A - \lambda I)c = 0.$$

For example 10, we have

$$\begin{bmatrix} 5 - \lambda & 1 \\ 2 & 4 - \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The two values for $\lambda$ are 6 and 3. Inserting these values in the above expression yields the following:

- For $\lambda = 6$, $-c_1 + c_2 = 0$ and $2c_1 - 2c_2 = 0$, or $c_1 = c_2$.
- For $\lambda = 3$, $2c_1 + c_2 = 0$ and $2c_1 + c_2 = 0$, or $c_1 = -\frac{1}{2}c_2$.

Note that neither pair determines the values of $c_1$ and $c_2$.

It was the reason $c'c = 1$ was specified at the outset.

So, for example, if $\lambda = 6$, any vector $c$ with equal elements will satisfy the expression above.

The additional equation $c'c = 1$, however, produces complete solutions for both vectors:

For $\lambda = 6$, $c = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.

For $\lambda = 3$, $c = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}$. 
These are the only vectors that satisfy both the expression above and $c'c = 1$.

**General Results for Eigenvalues and Eigenvectors**

A $K \times K$ symmetric matrix has $k$ distinct eigenvectors, $c_1, c_2, \ldots, c_K$.

The corresponding eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_K$, although real, need not be distinct.

The eigenvalues of a symmetric matrix are orthogonal. This implies that for every $i \neq j$, $c_i'c_j = 0$

It is convenient to collect the $K$ eigenvectors in a $K \times K$ matrix whose $i$th column is the $c_i$ corresponding to $\lambda_i$,

$$C = \begin{bmatrix} c_1 & c_2 & \cdots & c_K \end{bmatrix},$$

and the $K$ eigenvalues in the same order, in a diagonal matrix,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \cdots & \lambda_K \end{bmatrix}.$$

Then, the full set of equations

$$Ac_i = \lambda_i c_i$$

is contained in

$$AC = CA.$$

Since the vectors are orthogonal and $c_i'c_i = 1$, we have

$$C'C = \begin{bmatrix} c_1'c_1 & c_1'c_2 & \cdots & c_1'c_K \\ c_2'c_1 & c_2'c_2 & \cdots & c_2'c_K \\ & & \ddots & \\ c_K'c_1 & c_K'c_2 & \cdots & c_K'c_K \end{bmatrix} = I.$$

This implies that $C' = C^{-1}$.

Consequently, $CC' = CC^{-1} = I$ as well, so the rows as well as the columns of $C$ are orthogonal.

**Eigenvalues, Eigenvectors and Determinant of a Matrix.** Recalling how tedious the calculation of the determinant promised to be, we find that the following is particularly useful. Since
\[ C'AC = \Lambda, \]
\[ |C'AC| = |\Lambda|. \]

Using a number of algebraic properties of matrices, we have, for orthogonal matrix \( C \),
\[
|C'AC| = |C'| \cdot |A| \cdot |C| = |C'| \cdot |C| \cdot |A| = |C'C| \cdot |A|
\]
\[
= |I| \cdot |A| = 1 \cdot |A|
\]
\[
= |A|
\]

Since \( |\Lambda| \) is just the product of its diagonal elements, this implies the following.

**Theorem 2** *The determinant of a matrix equals the product of its eigenvalues.*

Notice that we get the expected result if any of these eigenvalues is zero.

Since the determinant is the product of the characteristic roots, it follows that a matrix is singular if and only if its determinant is zero and, in turn, if and only if it has at least one zero eigenvalue.

**Multivariate Maximization**

Let us now consider the next level of complexity of maximization problems.

Suppose that we have now two choice variables \( x_1 \) and \( x_2 \).

It will often be convenient to write these two variables as a vector \( \mathbf{x} = (x_1, x_2) \).

In this case, we write the maximization problem using the notation

\[
\max_{x_1, x_2} f(x_1, x_2)
\]

or, more generally, as

\[
\max_{\mathbf{x}} f(\mathbf{x})
\]

For maximizing or minimizing a function of several variables, the first order conditions are:

\[
\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = 0
\]

At the optimum, it must be true that no small change in any variable leads to an improvement in the function value.
In other words, at the optimal choice \( x^* \) the vector of partial derivatives must equal the zero vector.

In this case, given that we only have two choice variables, there are two necessary conditions:

\[
\frac{\partial f(x_1, x_2)}{\partial x_1} = 0
\]

\[
\frac{\partial f(x_1, x_2)}{\partial x_2} = 0.
\]

The second-order conditions for the two-choice variable problem are most easily expressed in terms of the matrix of second derivatives of the objective function.

This matrix, as we already know, takes the name of Hessian matrix, and takes the form

\[
H = \begin{bmatrix}
 f_{11} & f_{12} \\
 f_{21} & f_{22}
\end{bmatrix},
\]

where \( f_{ij} \) stands for \( \frac{\partial^2 f}{\partial x_i \partial x_j} \).

Calculus tells us that at the optimal choice \( x^* \), the Hessian matrix must be negative semidefinite.

This means that for any vector \( (h_1, h_2) \), we must satisfy

\[
\begin{bmatrix}
 h_1 \\
 h_2
\end{bmatrix}
\begin{bmatrix}
 f_{11} & f_{12} \\
 f_{21} & f_{22}
\end{bmatrix}
\begin{bmatrix}
 h_1 \\
 h_2
\end{bmatrix} \leq 0.
\]

More generally, let us think of \( h \) as a column vector and let \( h' \) be the transpose of \( h \).

Then, we can write the condition characterizing a negative semidefinite matrix as

\[ h'Hh \leq 0. \]

If we are examining a minimization problem rather than a maximization problem, the first-order condition is the same, but the second-order condition becomes the requirement that the Hessian matrix be positive semidefinite.

When we are dealing with a function \( y = f(x_1, x_2, \cdots, x_n) \), or \( y = f(\mathbf{x}) \), it is very useful to write the function in matrix form and solve the maximization problem using matrix operations.

For example, let \( y = \sum_{i=1}^{n} a_i x_i \).
We can rewrite this function as \( y = a'x = x'a \).

The partial derivative of this function is,

\[
\frac{\partial (a'x)}{\partial x} = a.
\]

Note, in particular, that \( \frac{\partial (a'x)}{\partial x} = a \), not \( a' \).

In a set of linear functions \( y = Ax \), each element \( y_i \) of \( y \) is

\[
y_i = a_i'x,
\]

where \( a_i \) is the \( i \)th row of \( A \).

Therefore, \( \frac{\partial y_i}{\partial x} = a_i' \), or the transpose of the \( i \)th row of \( A \), and

\[
\begin{bmatrix}
    \frac{\partial y_1}{\partial x} \\
    \frac{\partial y_2}{\partial x} \\
    \vdots \\
    \frac{\partial y_n}{\partial x}
\end{bmatrix} =
\begin{bmatrix}
a_1' \\
a_2' \\
\vdots \\
a_n'
\end{bmatrix}.
\]

Collecting all terms, we find that \( \frac{\partial Ax}{\partial x'} = A \), whereas the more familiar form will be

\[
\frac{\partial (Ax)}{\partial x} = A'.
\]

Now, let \( y = \sum_{i=1}^{n} \sum_{j=1}^{n} x_ix_ja_{ij} \). We can rewrite this expression as \( x'Ax \).

So, for example, suppose \( A = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} \), so that \( x'Ax = 1x_1^2 + 4x_2^2 + 6x_1x_2 \).

Then

\[
\frac{\partial x'Ax}{\partial x} =
\begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 8x_2 \end{bmatrix} =
\begin{bmatrix} 2 & 6 \\ 6 & 8 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2Ax,
\]

which is the general result when \( A \) is a symmetric matrix.

If \( A \) is not symmetric, then

\[
\frac{\partial (x'Ax)}{\partial x} = (A + A')x.
\]
Referring to the preceding double summation, we find that for each term, the coefficient on $a_{ij}$ is $x_i x_j$. Therefore

$$\frac{\partial (x'Ax)}{\partial a_{ij}} = x_i x_j.$$ 

The square matrix whose $ij$th element is $x_i x_j$ is $xx'$, so

$$\frac{\partial (x'Ax)}{\partial A} = xx'.$$

**Example 3** Consider the problem

$$\max_x f(x) = \sum_{i=1}^{n} a_i x_i - \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j b_{ij}$$

where, $a = \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 2 \\ 3 & 2 & 5 \end{bmatrix}$.

Our function can be written as

$$\max_x f(x) = a'x - x'Bx,$$

Using some now familiar results, we obtain

$$\frac{\partial f(x)}{\partial x} = a - 2Bx$$

$$= \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 4 \\ 6 & 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$ 

The solutions are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 4 \\ 6 & 4 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 11.25 \\ 1.75 \\ -7.25 \end{bmatrix}.$$ 

The sufficient condition is that

$$\frac{\partial^2 f(x)}{\partial x \partial x'} = -2B = \begin{bmatrix} -4 & -2 & -6 \\ -4 & -6 & -4 \\ -6 & -4 & -10 \end{bmatrix}$$

must be negative definite.
The three eigenvalues of this matrix are
\[ \lambda_1 = -15.746, \lambda_2 = -4, \text{ and } \lambda_3 = -0.25403. \]

Since all three eigenvalues are negative, the matrix is negative definite, as required.

Note that we needed to compute the eigenvalues of the Hessian to verify the sufficient condition.

For a general matrix of order larger than 2, this will normally require a computer. Suppose, however, that \( A \) is of the form
\[ A = B'B \]
where \( B \) is some known matrix. Then, we know that \( A \) will always be positive definite (assuming \( B \) has full rank).

Therefore, it will not be necessary to calculate the eigenvalues of \( A \) to verify the sufficient conditions.

Comparative Statics.

Suppose that \( f(x_1, x_2) \), and that both \( x_1 \) and \( x_2 \) are functions of \( a \) (i.e. \( x_1(a) \) and \( x_2(a) \)). Just as before, we may want to determine how the optimal choice changes as the parameter \( a \) changes.

We know that the optimal choices have to satisfy the first-order conditions
\[ \frac{\partial f(x_1(a), x_2(a), a)}{\partial x_1} = 0 \]
\[ \frac{\partial f(x_1(a), x_2(a), a)}{\partial x_2} = 0. \]

Differentiating these two expressions with respect to \( a \), we have
\[ \frac{\partial^2 f}{\partial x_1^2} \frac{\partial x_1}{\partial a} + \frac{\partial^2 f}{\partial x_1 \partial x_2} \frac{\partial x_2}{\partial a} + \frac{\partial^2 f}{\partial x_1 \partial a} = 0 \]
\[ \frac{\partial^2 f}{\partial x_1 \partial x_2} \frac{\partial x_1}{\partial a} + \frac{\partial^2 f}{\partial x_2^2} \frac{\partial x_2}{\partial a} + \frac{\partial^2 f}{\partial x_2 \partial a} = 0 \]

This is more conveniently written in matrix form as
\[ \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial a} \\ \frac{\partial x_2}{\partial a} \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 f}{\partial x_1 \partial x_2} \\ -\frac{\partial^2 f}{\partial x_2 \partial a} \end{bmatrix}. \]
If the matrix on the LHS of this expression is invertible, we can solve this system of equations to get

$$\begin{bmatrix}
\frac{\partial x_1}{\partial a} \\
\frac{\partial x_2}{\partial a}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2}
\end{bmatrix}^{-1} \begin{bmatrix}
-\frac{\partial^2 f}{\partial x_1 \partial a} \\
-\frac{\partial^2 f}{\partial x_2 \partial a}
\end{bmatrix}.$$  

Rather than invert the matrix, it is often easier to use Cramer’s rule to solve the system of equations for $\frac{\partial x_1}{\partial a}$ and $\frac{\partial x_2}{\partial a}$.

For example, if we want to solve for $\frac{\partial x_1}{\partial a}$, we can apply Cramer’s rule to express this derivative as the ratio of two determinants:

$$\frac{\partial x_1}{\partial a} = \frac{\begin{vmatrix}
-\frac{\partial^2 f}{\partial x_1 \partial a} \\
-\frac{\partial^2 f}{\partial x_2 \partial a}
\end{vmatrix}}{\begin{vmatrix}
\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2}
\end{vmatrix}}.$$

By the second-order condition for maximization, the matrix in the denominator of this expression is a negative semidefinite matrix.

Matrix algebra tells us that this matrix must have a positive determinant.

Therefore, the sign of $\frac{\partial x_1}{\partial a}$ is simply the sign of the determinant in the numerator.

**Example 4** Let $f(x_1, x_2, a_1, a_2) = u_1(x_1) + u_2(x_2) - a_1x_1 - a_2x_2$.

The first-order conditions for maximizing $f$ are

$$u_1'(x_1^*) - a_1 = 0$$
$$u_2'(x_2^*) - a_2 = 0$$

The second order condition is that the matrix

$$H = \begin{bmatrix}
u_1''(x_1^*) & 0 \\
0 & u_2''(x_2^*)
\end{bmatrix}$$

is negative semidefinite.

Since a negative semidefinite matrix must have diagonal terms that are less than or equal to zero, it follows that $u_1''(x_1^*) \leq 0$ and $u_2''(x_2^*) \leq 0$. 

17
The maximized value function is given by

\[ M(a_1, a_2) \equiv \max_{x_1, x_2} u_1(x_1) + u_2(x_2) - a_1 x_1 - a_2 x_2, \]

and a simple calculation using the envelope theorem shows that

\[ \frac{\partial M}{\partial a_1} = -x_1^* \]
\[ \frac{\partial M}{\partial a_2} = -x_2^* \]

The comparative statics calculation immediately above shows that

\[ \text{sign} \frac{\partial x_1}{\partial a_1} = \text{sign} \left| \begin{array}{cc} 1 & 0 \\ 0 & u_2''(x_2^*) \end{array} \right|. \]

Carrying out the calculation of the determinant,

\[ \text{sign} \frac{\partial x_1}{\partial a_1} \leq 0. \]

Note that we can determine how the choice variable responds to changes in the parameter without knowing anything about the explicit functional form of \( u_1 \) or \( u_2 \); we only have to know the structure of the objective function—in this case, that it is additively separable.

**Markov Processes**

We are now going to explore another important application of eigenvalues and eigenvectors: the solution of Markov processes.

A stochastic process is the counterpart to a deterministic process. In the simplest possible case (“discrete time”), a stochastic process amounts to a sequence of a finite number of outcomes.

When an outcome at any step in the sequence depends, at most, on the outcome of the preceding step and not on any other previous outcome, such a sequence is called a Markov chain or Markov process.

Suppose that kids, \( A, B, C \) are throwing a ball to each other. \( A \) always throws the ball to \( B \), and \( B \) always throws the ball to \( C \). However, \( C \) is just as likely to throw the ball to \( B \) as to \( A \). This is an example of a Markov process: only the immediate past matters.
Transition Matrix of a Markov Process.

The key elements of a Markov process are:

1. Each outcome belongs to a finite set \( \{ a_1, a_2, ..., a_n \} \) called the state space of the system; if the outcome at time \( n \) is \( a_i \), then we say the system is in state \( a_i \) at time \( n \) or at the \( n^{th} \) step.

2. The outcome at any step depends, at most on the outcome of the preceding step and not on any other previous outcome.

Accordingly, with each pair of states \((i, j)\), there is given the probability \( m_{ij} \) that the process will be in state \( i \) immediately after \( j \) occurs. Notice that these probabilities are written so that the first subscript indexes the next period and the second subscript indexes the current period.

These transition probabilities \( m_{ij} \) form the following square matrix:

\[
M = \begin{bmatrix}
  m_{11} & m_{12} & \cdots & m_{1n} \\
  m_{21} & m_{22} & \cdots & m_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{n1} & m_{n2} & \cdots & m_{nn}
\end{bmatrix}
\]

This matrix \( M \) is called the transition matrix of the Markov process. Observe that with each state \( j \) there corresponds the \( j^{th} \) column of the transition matrix \( M \). Moreover, if the system is in state \( j \), then this column represents all the possible outcomes in the next period and so it is a probability vector.

A vector \( \mathbf{q} = [q_1, q_2, ..., q_n] \) is called a probability vector if its entries are nonnegative and their sum is 1, that is, if:

1. Each \( q_i \geq 0 \),
2. \( q_1 + q_2 + ... + q_n = 1 \).

A square matrix \( \mathbf{P} = [p_{ij}] \) is called a stochastic matrix if each column (or row) of \( \mathbf{P} \) is a probability vector. Therefore, the transition matrix \( M \) of a Markov process is a stochastic matrix.

**Example 5** Three kids, Ann (A), Bill (B), and Charlie (C), are throwing a ball to each other. Ann always throws the ball to Bill, and Bill always throws the ball to Charlie. However, Charlie is just as likely to throw the ball to Bill as to Ann.

The ball throwing is a Markov process with the following transition matrix:
\[ M = \begin{bmatrix}
0 & 0 & \frac{1}{2} \\
1 & 0 & \frac{1}{2} \\
0 & 1 & 0
\end{bmatrix}. \]

The first column of the matrix corresponds to the fact that Ann always throws the ball to Bill. The second column of the matrix corresponds to the fact that Bill always throws the ball to Charlie. The last column of the matrix corresponds to the fact that Charlie always throws the ball to Ann or Bill with equal probability (and does not throw the ball to himself).

We now define an important class of stochastic matrices: a stochastic matrix \( P \) is said to be regular if all entries of some power \( P^m \) of \( P \) are positive.

Consider the following matrix \( A = \begin{bmatrix}
0 & 1 \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}. \) This matrix is regular because all entries in \( A^2 \) are positive:

\[
A^2 = \begin{bmatrix}
0 & 1 \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix} \begin{bmatrix}
0 & 1 \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{4} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{4}
\end{bmatrix}
\]

Some fundamental properties of regular stochastic matrices are contained in the following theorem:

**Theorem 6** Let \( P \) be a regular stochastic matrix. Then:

1. \( P \) has a unique fixed probability vector \( t \), and the components of \( t \) are all positive.
2. If \( q \) is any probability vector, then the sequence of vectors \( q, Pq, P^2q, P^3q \ldots \) approaches the fixed point \( t \).

State Distributions

Consider a Markov process with transition matrix \( M \). Suppose that the initial state distribution \( q_0 \) (at time=0) is given. Then the subsequent state distributions can be obtained by multiplying the preceding state distribution by the transition matrix \( M \). Namely,

\[ Mq_0 = q_1, Mq_1 = q_2, Mq_2 = q_3, \ldots \]

Accordingly,

\[ q_2 = Mq_1 = (MM)q_0 = M^2q_0 \]

and so on. We state this result formally:

Suppose an initial state distribution \( q_0 \) is given. Then, for \( k = 1, 2, \ldots, \)

\[ q_k = Mq_{k-1} = M^k q_0 \]

20
Example 7 Consider the Markov chain in the preceding example, with transition matrix $M$.

Suppose Charlie is the first person with the ball, that is suppose $q_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is the initial probability distribution.

Then,
\[
q_1 = Mq_0 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix},
\]
\[
q_2 = Mq_1 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix},
\]
\[
q_3 = Mq_2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}.
\]

Therefore, after 3 throws, the probability that Ann has the ball is $\frac{1}{4}$, that Bill has the ball is $\frac{1}{2}$, and that Charlie has the ball is $\frac{1}{4}$.

Regular Markov Processes and Stationary State Distributions

A Markov chain is said to be regular if its transition matrix $M$ is regular. Recall that if $M$ is regular, then $M$ has a unique fixed probability vector $t$ and, for any probability vector $q$, the sequence of vectors $q, Mq, M^2q, M^3q...$ approaches the fixed point $t$.

Theorem 8 Suppose the transition matrix $M$ of a Markov chain is regular. Then, in the long run, the probability that any state $a_i$ occurs is approximately equal to the component $t_i$ of the unique fixed probability vector $t$ of $M$.

In other words, the effect of the initial state distribution in a regular Markov process wears off as the number of steps increases.

Every sequence of state distributions approaches the fixed probability vector $t$ of $M$, which is called the stationary distribution of the Markov chain.

Example 9 Consider the Markov chain in the preceding example, where Ann, Bill, and Charlie throw a ball to each other with the following transition matrix:
\[
M = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}.
\]
To find its unique fixed probability vector $t$ we need to find a vector $t$ such that $Mt = t$. The vector can be represented in the form \[
\begin{bmatrix}
x \\
y \\
1 - x - y
\end{bmatrix}.
\]
Accordingly, we form the following matrix equation:

\[
\begin{bmatrix}
0 & 0 & \frac{1}{2} \\
1 & 0 & \frac{1}{2} \\
0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
1 - x - y
\end{bmatrix} =
\begin{bmatrix}
x \\
y \\
1 - x - y
\end{bmatrix}
\]

This yields the following system of linear equations:

\[
\frac{1}{2} - \frac{1}{2}x - \frac{1}{2}y = x
\]
\[
x + \frac{1}{2} - \frac{1}{2}x - \frac{1}{2}y = y
\]
\[
1 - x - y = y
\]

or

\[
3x + y = 1
\]
\[
x + x - 3y = -1
\]
\[
x + 2y = 1
\]

And solving for $x$ and $y$, we obtain $x = \frac{1}{5}$ and $y = \frac{2}{5}$. Therefore, $t = \begin{bmatrix}
\frac{1}{5} \\
\frac{2}{5}
\end{bmatrix}$.

So, we can conclude that in the long run, Ann will be thrown the ball 20 percent of the time, and Bill and Charlie will be thrown the ball 40 percent of the time.

Using Eigenvalues and Eigenvectors

Notice that to find the stationary distribution of the Markov chain we constructed a system of linear equations. As usual, another way of solving the problem is to use matrix algebra.

Consider the following Markov process. A man either takes a bus or drives his car to work each day. Suppose he never takes the bus 2 days in a row; but if he drives to work, then the next day he is just as likely to drive again as he is to take the bus.

This stochastic process is a Markov chain because the outcome on any day depends only on what happened the preceding day.

In this case, the state space is \{bus, drive\} and the transition matrix $M$ follows:
The first column of the matrix corresponds to the fact that the man never takes the bus 2 days in a row, and so he definitively will drive the day after he takes the bus. The second column of the matrix corresponds to the fact that the day after he drives he will drive or take the bus with equal probability.

To find the stationary distribution of the Markov process (i.e. what would happen after \( n \) periods, or the outcome at period \( n+1 \)) we seek a probability vector \( \mathbf{x}(n+1) = \begin{bmatrix} x_1(n+1) \\ x_2(n+1) \end{bmatrix} \) such that \( \mathbf{M}\mathbf{x}(n) = \mathbf{x}(n+1) \). Accordingly, we form the following matrix equation:

\[
\begin{bmatrix} x_1(n+1) \\ x_2(n+1) \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}
\]

Notice that if we subtract a 1 from each diagonal entry of the matrix \( \mathbf{M} \), then each column of the transformed matrix adds up to 0:

\[
\mathbf{M} - \mathbf{1} \mathbf{I} = \begin{bmatrix} -1 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}
\]

Recall that if the columns of a square matrix add up to \((0,...,0)\), the rows are linearly dependent and the matrix must be singular. It follows that \( \lambda = 1 \) is an eigenvalue of the matrix \( \mathbf{M} \).

Now, we can find the other eigenvalue using the trace of the matrix.

The trace of a square matrix is the sum of its diagonal entries. In this case, the trace of the matrix \( \mathbf{M} = 0 + \frac{1}{2} = \frac{1}{2} \).

As the following theorem indicates, one can learn something about eigenvalues of a matrix just by examining the trace of the matrix.

**Theorem 10** Let \( \mathbf{A} \) be a \( k \times k \) matrix with eigenvalues \( \lambda_1, ..., \lambda_k \). Then:

\[ \lambda_1 + \lambda_2 + ... + \lambda_k = \text{trace of } \mathbf{A}. \]

In the case of the transition matrix \( \mathbf{M} \) under consideration, we already know that one of the eigenvalues \( \lambda_1 = 1 \); therefore, we can conclude that the other eigenvalue is

\[ \lambda_2 = \frac{1}{2} - 1 = -\frac{1}{2}. \]

To compute the corresponding eigenvectors, we solve

\[
\begin{bmatrix} 0 - \lambda_1 & \frac{1}{2} \\ 1 & \frac{1}{2} - \lambda_1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},
\]
and
\[
\begin{bmatrix}
0 - \lambda_2 & \frac{1}{2} \\
1 & \frac{1}{2} - \lambda_2
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2
\end{bmatrix}
= \begin{bmatrix}
0 \\
0
\end{bmatrix},
\]

The two values for \( \lambda \) are 6 and 3. Inserting these values in the above expression yields the following:

- For \( \lambda_1 = 1 \), \(-c_1 + \frac{1}{2}c_2 = 0 \) and \( c_1 - \frac{1}{2}c_2 = 0 \), or \( c_1 = 2c_2 \).
- For \( \lambda_2 = -\frac{1}{2} \), \( \frac{1}{2}c_1 + \frac{1}{2}c_2 = 0 \) and \( c_1 + c_2 = 0 \), or \( c_1 = -c_2 \).

Choosing the “simplest” of eigenvectors, the solution to this system of equations is
\[
\begin{bmatrix}
x_1(n+1) \\
x_2(n+1)
\end{bmatrix}
= \alpha \begin{bmatrix}
1 \\
2
\end{bmatrix} \cdot 1^n + \beta \begin{bmatrix}
1 \\
-1
\end{bmatrix} \cdot (-\frac{1}{2})^n
\]

Since \( 1^n = 1 \) and \((-\frac{1}{2})^n \to 0 \) as \( n \to \infty \), in the long-run, the solution of the Markov process tends to \( t = \alpha \begin{bmatrix}
1 \\
2
\end{bmatrix} \). Since \( t \) should be a probability vector whose components sum to 1, we can set \( \alpha \) to be the reciprocal \( \frac{1}{3} \) of the sum of the components of \( t \).

We conclude that the solution to the system of equations tends to \( \begin{bmatrix}
\frac{1}{3}
\end{bmatrix} \) as \( n \to \infty \).

Therefore, in the long run, the man will take the bus to work 33 percent of the time, and drive to work the other 66 percent of the time.

There are a number of general principles about Markov processes that the analysis of this examples illustrates:

1. \( \lambda_1 = 1 \) is always an eigenvalue of the underlying Markov matrix.
2. For a \( 2 \times 2 \) problem, the trace, which equals the sum of the eigenvalues, lies between 0 and 2. Therefore, the second eigenvalue \( \lambda_2 \) lies between -1 and +1.
3. The general solution is of the form presented above: since \( \lambda_2^n \to 0 \) as \( n \to \infty \), every solution tends to \( v \), where \( v \) is written as \( t \) divided by the sum of its components.
4. For a \( 2 \times 2 \) problem with a stochastic matrix of the form \( M = \begin{bmatrix}
1 - a & b \\
a & 1 - b
\end{bmatrix} \), the vector \( v = \begin{bmatrix}
b \\
a
\end{bmatrix} \) is a fixed point of \( M \) (and it consist of the non-diagonal elements of \( M \)).

Besides the example presented above, one more general principle regarding Markov processes is worth noting
- If each $a_{ii} < 1$, then the elements of $(A - 1I)$ have the sign pattern $= \begin{bmatrix} - & + \\ + & - \end{bmatrix}$.

Therefore eigenvalue $\lambda_1 = 1$ has an eigenvector with all positive entries. This eigenvector can be made a probability vector by dividing each of its components by their sum.

In fact, most of these properties hold for a large class of Markov processes – namely, processes in which the transition matrix is regular.

For instance, consider again the Markov chain in example of the three kids throwing the ball to each other. The data of this problem lead to the following Markov system:

$$
\begin{bmatrix}
    x_1(n + 1) \\
    x_2(n + 1) \\
    x_3(n + 1)
\end{bmatrix} =
\begin{bmatrix}
    0 & 0 & 1/2 \\
    1 & 0 & 1/2 \\
    0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
    x_1(n) \\
    x_2(n) \\
    x_3(n)
\end{bmatrix}
$$

We can use the fact that $\lambda_1 = 1$ is an eigenvalue of the underlying Markov matrix to calculate the corresponding eigenvector. In this case, we solve

$$
\begin{bmatrix}
    0 - \lambda_1 & 0 & 1/2 \\
    1 & 0 - \lambda_1 & 1/2 \\
    0 & 1 & 0 - \lambda_1
\end{bmatrix}
\begin{bmatrix}
    c_1 \\
    c_2 \\
    c_3
\end{bmatrix} =
\begin{bmatrix}
    0 \\
    0 \\
    0
\end{bmatrix},
$$

or

$$
\begin{bmatrix}
    -1 & 0 & 1/2 \\
    1 & -1 & 1/2 \\
    0 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
    c_1 \\
    c_2 \\
    c_3
\end{bmatrix} =
\begin{bmatrix}
    0 \\
    0 \\
    0
\end{bmatrix}.
$$

Therefore, the above expression yields the following:

$$-c_1 + \frac{1}{2}c_3 = 0$$

$$c_1 - c_2 + \frac{1}{2}c_3 = 0$$

$$c_2 - c_3 = 0$$

And solving for $c_1$, $c_2$ and $c_3$, we obtain $c_1 = 1$, $c_2 = 2$, and $c_3 = 2$. Therefore, $t = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ is a fixed point of $M$.

Because $\lambda_2^n \to 0$ and $\lambda_3^n \to 0$ as $n \to \infty$, every solution tends to $v$, where $v$ is written as $t$ divided by the sum of its components. In this case, as we already know, this means that the solution to the system of equations tends to $v = \begin{bmatrix} 1 \\ 5/2 \\ 5/2 \\ 5/2 \end{bmatrix}$ as $n \to \infty$. 

25