Overview of Accounting Choice, Linear Algebra, and Regression

Understanding the role of accounting involves understanding accounting choice. Accounting choice is the design of accounting systems to (hopefully) enhance productive efficiency and social welfare. The study of accounting choice involves a few key ingredients: a productive environment or technology to provide context, a rich but restrictive (usually linear) structure typically involving aggregation, and uncertainty.

**Linear algebra: subspaces and projections**

Accounting is a large system of linear equations. Accordingly, the systematic study of accounting complements and is complemented by the study of linear algebra. This section of notes provides an overview of fundamental results from linear algebra.\(^1\)

Large systems of equations frequently involve one of two issues: (i) there are many, many consistent solutions to the system, or (ii) there is no solution that satisfies the system of equations.\(^2\) The former situation usually means that there are fewer (linearly independent) equations than parameters to be determined (or estimated). The latter situation usually means that there are many equations and few parameters to be determined (or estimated). When there is too little data and accordingly many consistent solutions, the analyst is free to choose some variables (called free variables) so that the other variable values become identifiable. When there is too much data and no consistent solution, the analyst looks for the nearest distance (least squares) solution.

The **fundamental theorem of linear algebra** is an indispensable aid for gaining a deeper understanding of these two situations. The fundamental theorem of linear algebra says that any matrix (rectangular array of elements) is comprised of four fundamental subspaces: the rowspace, columnspace, nullspace, and left nullspace. The **rowspace** is

---

\(^1\) A brief review of matrix operations (addition, multiplication, and inverses) appears at the end of this section.

\(^2\) Occasionally there exists a unique solution.
the subspace mapped out via linear combinations of the rows. Accordingly, a **basis** is any complete set of linearly independent rows and the number of linearly independent rows is the **dimension** of the rowspace. The **columnspace** is the subspace mapped out via linear combinations of the columns. A **basis** is any complete set of linearly independent columns and the number of linearly independent columns is the **dimension** of the columnspace. The fundamental theorem states that the dimension of the rowspace and the dimension of the columnspace are equal for any matrix; this is also referred to as the **rank** \( r \) of the matrix. Hence, all matrices have the same number of linearly independent rows and columns.

The second part of the fundamental theorem deals with orthogonal subspaces. Two matrices (or vectors) are **orthogonal** (in Euclidean geometry this is seen as perpendicular) iff their product is zero. Since its product with the rows of a matrix yields zeroes, the subspace orthogonal to the rowspace is called the **nullspace**. For (m rows x n columns) matrix A, this means that \( AN^T = 0 \) where \( N \) is the \( (n-r \times n) \) nullspace of A; \( N^T \) means **transpose** of \( N \) or the rows are switched to become the columns of the matrix and vice versa, so \( N^T \) is \( n \) rows and \( n-r \) columns. The dimension of the nullspace equals \( n-r \) so that the dimensions of the rowspace \( (r) \) and nullspace \( (n-r) \) add to the number of columns \( (n) \). Analogously, the subspace orthogonal to the columnspace is the **left nullspace**. Again for \( (m \times n) \) matrix A, \( LA = 0 \) where \( L \) is the \( (m-r \times m) \) left nullspace of A. The dimension of the left nullspace equals \( m-r \) so that the dimensions of the columnspace \( (r) \) and left nullspace \( (m-r) \) add to the number of rows \( (m) \).

Now return to the two situations raised earlier. Suppose, generically, the system of equations is represented by \( Ay = x \), \( y \) is an \( n \)-length vector, and \( x \) is an \( m \)-length vector, and the number of linearly independent rows is less than the length of \( y \) \( (r < n) \). This is

---

3 **Linear combination** is the weighted sum of the target attribute. Hence, a linear combination of rows is the weighted sum of rows.

4 **Linear independence** means that the set is comprised of attributes (e.g., rows or columns) that cannot be formed from linear combinations of its remaining attributes (rows or columns). For example, the set of rows \([1\ 1],[0\ 1]\), and \([1\ 0]\) is not linearly independent since the sum of the latter two equals \([1\ 1]\) but the set of rows \([1\ 0\ 0]\), \([0\ 1\ 0]\), and \([0\ 0\ 1]\) is linearly independent.
the first situation in which there exist many consistent solutions. One can think of the
entire set of consistent solutions as \( y = y^R + y^N \) where \( y^R \) is the linear combination of the
rows of \( A \) consistent with \( Ay^R = x \) and \( y^N = N^T k \) where \( k \) is a \((n-r)\)-length vector of free
variables or weights on the basis vectors of the nullspace. Notice that since \( AN^T = 0 \), the
null component of \( y \), \( y^N = N^T k \), when multiplied by \( A \) always yields zero. Accordingly,
one is free to choose the elements of \( k \) to be any values (hence, they’re called free
variables) and the solution is still consistent with the system of equations \( Ay = x \)
\((Ay^R + y^N = Ay^R + 0 = x)\). From the fundamental theorem, the key is recognizing that
the null component gives rise to the freedom in the solution.

In the alternative second situation, suppose \( Ay = x \) involves many more rows of \( A \) than
the length of \( y \) (\( m > n \)) and, for simplicity the columns of \( A \) are linearly independent (\( r = n \)). Usually no consistent solution \( y \) exists. The analyst forms a solution from the linear
combination of the columns of \( A \) that is the nearest distance to \( x \). This is referred to as a
projection. While no solution exists to \( Ay = x \), a solution always exists (when \( r = n \)) if
both sides of the equations are multiplied by \( A^T \), \( A^T Ay = A^T x \). These equations are called
the normal equations, and can be solved by Gaussian elimination. For convenience, the
(least squares) solution for \( y \) is usually written as \( y = (A^T A)^{-1} A^T x \). Now the projection is
\( Ay = A(A^T A)^{-1} A^T x \) and the important leading matrix, \( P = A(A^T A)^{-1} A^T \), is called the
projection matrix. The projection matrix is symmetric (equal to its transpose) and
idempotent (its product with itself is the original matrix, \( PP = P \)). The intuition for the
idempotent property is if the linear combination of the columns nearest to \( x \) is already
determined and we try to project onto the columns of \( A \) again, the solution doesn’t
change. How could it? We were already where we wanted to be.

One additional property should be mentioned before moving on. The left nullspace is the
orthogonal complement to the columnspace and together the two subspaces span \( m \)-
space, that is, any \( m \)-length vector can be represented by a linear combination of bases
for the columnspace and left nullspace. Thus, one could also project \( x \) onto the left

---

\(^5\) The notation \( A^{-1} \) refers to the inverse of \( A \). The inverse is the matrix when multiplied
by \( A \) yields the identity matrix \( I \), \( A^{-1} A = AA^{-1} = I \).
nullspace and subtract this projection from $x$ to find the projection onto the columnspace. That is, since $P_x + P_Lx = Ix = x$, $P_Ix = L(L^TL)^{-1}L^Tx = (I - P)x$ and $(I - P_L)x = x - P_Lx = Px$, where $I$ is the identity matrix. Occasionally, like with accounting structure, this relationship can reduce the size of the computational exercise (more on this later).

**Appendix: Basic matrix operations**

Scalar operations with matrices: Multiplication of a matrix $A$ by a scalar $b$ simply involves multiplication of each element of $A$ by $b$.

Example: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times 7 = \begin{bmatrix} 1 \times 7 & 2 \times 7 & 3 \times 7 \\ 4 \times 7 & 5 \times 7 & 6 \times 7 \end{bmatrix} = \begin{bmatrix} 7 & 14 & 21 \\ 28 & 35 & 42 \end{bmatrix}$.

Addition: Two matrices can be added if each has the same number of rows and columns as the matrix sum is the sum of row-column elements.

Example: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 + 7 & 2 + 8 & 3 + 9 \\ 4 + 10 & 5 + 11 & 6 + 12 \end{bmatrix} = \begin{bmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix}$.

Multiplication: Two matrices can be multiplied if the number of columns of the first matrix equals the number of rows in the second matrix (order matters) since the row $i$, column $j$ element of the product matrix is the sum of the product of the $(i,k)$ element of the first matrix and $(k,j)$ element of the second matrix. The product matrix has the same number of rows as the leading matrix and the same number of columns as the trailing matrix.

Example: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 1 \times 7 + 2 \times 9 & 1 \times 8 + 2 \times 10 \\ 3 \times 7 + 4 \times 9 & 3 \times 8 + 4 \times 10 \\ 5 \times 7 + 6 \times 9 & 5 \times 8 + 6 \times 10 \end{bmatrix} = \begin{bmatrix} 25 & 28 \\ 57 & 64 \\ 89 & 100 \end{bmatrix}$.

Inverses: The key idea related to inverses is the existence of an identity matrix. An identity matrix is the matrix $I$ when multiplied by any other matrix, say $A$, produces a

---

6 The identity matrix is a matrix $I$ that multiplies by any matrix $A$ to yield $A$ once again, $IA = A$. Hence, the identity matrix is a square, diagonal matrix (all off-diagonal elements are zero) of ones along the principal (upper left to lower right) diagonal.
product equal to $A$. The matrix $I$ is a square, diagonal matrix with diagonal elements equal to one. Now, the inverse of $A$, say $A^{-1}$, produces $AA^{-1} = A^{-1}A = I$. Matrix inverse (when it exists) accommodates matrix division in the sense that scalar division is equivalent to multiplication by the reciprocal (or scalar inverse).

**Accounting example**

Consider an accounting example. The accounting platform is common knowledge; that is, we know the accounts reported in the financial statements and the nature but not the amount of the transactions that produced the financial statements. Indeed the task is to determine the transactions amounts that generated the observed financial statements. Consider the following example.

<table>
<thead>
<tr>
<th>Balance Sheet</th>
<th>Ending Balance</th>
<th>Beginning Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Receivables</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Inventory</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Plant</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Total Assets</td>
<td>33</td>
<td>29</td>
</tr>
<tr>
<td>Payables</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Owners' Equity</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>Total Liab. &amp; Equity</td>
<td>33</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
</tr>
<tr>
<td>CGS</td>
</tr>
<tr>
<td>G&amp;A</td>
</tr>
<tr>
<td>Income</td>
</tr>
<tr>
<td>Activity</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Collections of receivables</td>
</tr>
<tr>
<td>Cash purchase of plant</td>
</tr>
<tr>
<td>Payment of payables</td>
</tr>
<tr>
<td>Bad debt expense</td>
</tr>
<tr>
<td>Credit sales</td>
</tr>
<tr>
<td>Depreciation - period cost</td>
</tr>
<tr>
<td>Recognition of CGS</td>
</tr>
<tr>
<td>Accrued expenses</td>
</tr>
<tr>
<td>Purchase inventory on credit</td>
</tr>
<tr>
<td>Depreciation - product cost</td>
</tr>
</tbody>
</table>

Financial Statements and allowable journal entries

The algebraic representation of the accounting system is an incidence matrix, denoted $A$, which has a row for every account and a column for every journal entry. An incidence matrix, a matrix in which every column consists of zeros, a positive one, and a negative one, is well-suited to capture the mechanics of the double entry system; the one in the column specifies the account debited and the negative one the credit account. Column one for the example will have a one in the cash row and a minus one in the receivables row. The $A$ matrix for the example is presented below.
The given information can also be represented geometrically in a directed graph. In the figure below the nodes are accounts, and the arcs are journal entries.

The arrowhead denotes the account debited, the tail the credit. The y's are unknown. All the information in the algebraic representation for a given x and A is contained in the directed graph representation.
The complementary nature of these two representations is helpful for identifying the subspaces for the incidence matrix, A. A basis for the rowspace of A is any seven (of the eight) rows; recall accounting is redundant in that Assets = Equities implying that there are seven linearly independent rows. If there are seven linearly independent rows then there are also seven linearly independent columns.

Identifying a basis for the column space is a bit more challenging but yet still made relatively simple by the structure of accounting. We know that a basis must include columns 5 and 7 plus five more. The geometric representation greatly simplifies the task. Any spanning tree, a directed graph in which every node is connected via arcs and has no loops, is a basis for the column space. One spanning tree is exhibited below.\(^7\)

\[
\begin{array}{c}
\text{Receivables} \\
1 \\
y1 \\
Payables \\
-3 \\
y4 \\
\text{Inventory} \\
-1 \\
y7 \\
\text{CGS} \\
3 \\
y8 \\
\text{G & A Exp} \\
3 \\
y5 \\
\text{Sales} \\
-7 \\
y10 \\
\text{Plant} \\
1 \\
\text{Cash} \\
3 \\
y2 \\
\end{array}
\]

Spanning tree with transactions 3, 6, and 9 deleted

This particular solution is \(y^p = [6, 3, 0, 0, 7, 0, 3, 3, 0, 2]\) and of course it is a consistent solution for the problem posed; that is, \(Ay^p = x\).

\(^7\) The **matrix tree theorem** identifies the number of different spanning trees that exist for an incidence matrix as \(|A_0A_0^T|\), the determinant of the symmetric matrix determined by A excluding one row multiplied by the transpose, or alternatively, \(|NN^T|\). For this example, there are 36 spanning trees.
A basis for the nullspace is any set of linearly independent loops in the graph. For instance, $N = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & -1 \end{bmatrix}$ is a basis for the nullspace.

Euler’s formula says that for any connected graph the number of linearly independent loops is the number of arcs minus the number of nodes plus one ($n - m + 1$ loops).

Likewise, a basis for the left nullspace is the balancing vector (recall the balancing property of accounting, Assets = Equities and Debits = Credits). This is a vector of ones and usually denoted $\iota$ (iota). Hence, the subspaces for a double entry accounting matrix are straightforward to identify.

We have already identified one of the many consistent solutions (which we called $y^p$) for $Ay = x$ from the spanning tree above. The next question is how to represent compactly all possible solutions. Recall that the null component of transactions when multiplied by $A$ is equal to zero. Hence, all consistent solutions can be represented by $y = y^p + N^Tk$, where $k$ is a $(3 \times 1)$-vector of free variables $[k_1, k_2, k_3]$.

Of course, there are many (an arbitrarily large number) of consistent $y^p$’s. Is there a unique solution? A unique solution can be defined by the shortest length vector that is consistent with $Ay = x$. This is the row component of transactions and is found via a projection of (any) $y^p$ onto the rows of $A_0$; that is $y^R = A_0^T(A_0A_0^T)^{-1}A_0y^p$. A computationally simpler, yet equivalent, solution involves projecting $y^p$ onto the nullspace of $A$ and subtracting from $y^p$; that is $y^R = (I - N^T(NN^T)^{-1}N)y^p$. This completes the illustration of the fundamental algebraic relations for the (linear) accounting system.