Abstract

Equations of the form
\[ \sum_{i=1}^{n} x_i^k = r_k \quad \text{for } k = 1, 2, \ldots, n \]

have solutions \( x_i \) that are roots of polynomials whose coefficients are functions of the \( r_k \). When there are more equations than unknowns, and/or when \( k \) takes (positive integer) values other than the sequence 1 through \( n \), the equations can be reduced to the above form and solved by means of the appropriate polynomial.
EQUATIONS IN SUMS OF POWERS

S. R. Searle
August, 1966
Biometrics Unit, Cornell University, Ithaca, N.Y.

Identities

We first recall certain identities among sums of powers and sums of products of n values \(x_1, x_2, \ldots, x_n\). To do so, \(s_k\) is defined as the sum of the \(k'\)th powers:

\[
s_k = \sum_{i=1}^{n} x_i^k = x_1^k + x_2^k + \cdots + x_n^k;
\]

and \(p_t\) is defined as the sum of products of the \(x\)'s taken \(t\) at a time; for example,

\[
p_2 = \sum_{i \neq j} x_i x_j \quad \text{and} \quad p_3 = \sum_{i \neq j \neq k} x_i x_j x_k.
\]

In particular, when there are 4 \(x\)-values, i.e., \(n = 4\),

\[
s_1 = x_1 + x_2 + x_3 + x_4
\]

\[
p_1 = x_1 + x_2 + x_3 + x_4
\]

\[
p_2 = x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4
\]

\[
p_3 = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4
\]

and

\[
p_4 = x_1 x_2 x_3 x_4
\]

These definitions clearly imply that \(p_1\) and \(s_1\) are the same. Indeed, every \(p\) can be expressed in terms of \(s\)'s. Thus

\[
p_1 = s_1
\]

\[
p_2 = \frac{1}{2} (s_1^2 - s_2)
\]

\[
p_3 = \frac{1}{6} (s_1^3 - 3s_1 s_2 + 2s_3)
\]

\[
p_4 = \frac{1}{24} (s_1^4 + 8s_1 s_3 - 6s_1 s_2 + 3s_2^2 - 6s_4)
\]
and so on. (2) is obvious, (3) is well-known and (4) is easily verified by expanding \( (\sum x_i)(\sum x_i x_j) = s_1 p_2 \) making use of (3). In like fashion (5) can be established by expanding such products as \( s_1 p_3 \), \( s_2^2 \) or \( s_1^4 \). These results are very familiar in the context of the theory of equations. There, the \( p \)-terms are known as symmetric functions, and the above relationships are given in determinantal form (Turnbull, 1947) as

\[
p_1 = s_1 \\
p_2 = \frac{1}{2!} \begin{vmatrix} s_1 & 1 \\ s_2 & s_1 \end{vmatrix} \\
p_3 = \frac{1}{3!} \begin{vmatrix} s_1 & 1 & 0 \\ s_2 & s_1 & 2 \\ s_3 & s_2 & s_1 \end{vmatrix} \\
p_4 = \frac{1}{4!} \begin{vmatrix} s_1 & 1 & 0 & 0 \\ s_2 & s_1 & 2 & 0 \\ s_3 & s_2 & s_1 & 3 \\ s_4 & s_3 & s_2 & s_1 \end{vmatrix}
\]

and, in general

\[
p_t = \frac{1}{t!} \begin{vmatrix} s_1 & 1 & 0 & \cdots & \cdots & 0 \\ s_2 & s_1 & 2 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots & \vdots \\ s_i & s_{i-1} & \cdots & s_1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots & \vdots \\ s_t & s_{t-1} & \cdots & \cdots & \cdots & s_1 \end{vmatrix} \quad - - - (6)
\]
Note that in these relationships \( p_t \) does not involve any \( s_i \) for \( i > t \). This is, of course, clear from the definition of \( s_i \) and \( p_t \); and it leads to analogous expressions for the \( s_i \)'s in terms of the \( p_i \)'s - of which no use will be made here. Note too, that by definition

\[
p_t = 0 \quad \text{for all } t > n. \quad - - - (7)
\]

This is so because the definition of \( p_t \) must, for \( t > n \), imply the existence of at least \( t - r \) zeros associated with the \( n \) \( x \)'s; hence \( p_t \) is zero. This result is easily illustrated by considering \( p_3 \) for \( n = 2 \): from (4)

\[
6p_3 = (x_1 + x_2)^3 - 3(x_1 + x_2)(x_1^2 + x_2^2) + 2(x_1^3 + x_2^3)
= (x_1 + x_2)(-2)(x_1^2 - x_1x_2 + x_2^2) + 2(x_1^3 + x_2^3)
= 0.
\]

Solving equations

With the \( p_i \)'s as defined above it is well-known (loc. cit.) that the equation in \( y \) whose roots are \( x_1, x_2, \ldots, x_n \) can be written as

\[
y^n - p_1y^{n-1} + p_2y^{n-2} + \cdots + (-1)^{i}p_iy^{n-i} + \cdots + (-1)^{n-1}p_{n-1}y + (-1)^n p_n = 0.
\]

This is readily verified by expanding

\[
(y - x_1)(y - x_2)(y - x_3) \cdots (y - x_n) = 0.
\]

Suppose now, that we have \( n \) equations

\[
s_i = r_i \quad \text{for } i = 1,2,\ldots,n \quad - - - (9)
\]

to be solved for the \( n \) \( x \)-values implicit in the \( s_i \)-terms, the \( r_i \) being given numerical values. If \( p \)-values are calculated from (6) making use of (9), and if the resulting \( p_i \)'s are then used in the polynomial (8), the roots of this polynomial will be solutions to (9). For clearly, the roots of the polynomial will have the sums of powers given in (9) and hence they satisfy (9). Thus when \( n = 3 \) the given equations would be
\[ x_1 + x_2 + x_3 = r_1 \]
\[ x_1^2 + x_2^2 + x_3^2 = r_2 \]
\[ x_1^3 + x_2^3 + x_3^3 = r_3 . \]

Using \( r \)'s for \( s \)'s in (2), (3) and (4) we have

\[ p_1 = r_1 \]
\[ p_2 = \frac{1}{2} (r_1^2 - r_2) \]
\[ p_3 = \frac{1}{6} (r_1^3 - 3r_1r_2 + 2r_3) \]

and these \( p \)-values put into (8) give

\[ y^3 - r_1y^2 + \frac{1}{2} (r_1^2 - r_2)y - \frac{1}{6} (r_1^3 - 3r_1r_2 + 2r_3) = 0 . \]

The roots of this polynomial are the solutions to (10). Of course, because (10)
[and, in general, (1)] is symmetric in the \( x \)'s no distinction is made as to which of the solutions is \( x_1 \), \( x_2 \) or \( x_3 \). As an example, suppose equations (10) are

\[ x_1 + x_2 + x_3 = 7 , \]
\[ x_1^2 + x_2^2 + x_3^2 = 21 , \]
\[ x_1^3 + x_2^3 + x_3^3 = 73 . \]

Then (11) is

\[ y^3 - 7y^2 + \frac{1}{2} (49 - 21)y - \frac{1}{6} (343 - 441 + 146) = 0 \]

i.e.

\[ y^3 - 7y^2 + 14y - 8 = 0 , \]

or

\[ (y - 1)(y^2 - 6y + 8) = 0 , \]

which is satisfied by \( y = 1, 2 \) and 4. Hence \( 1, 2, 4 \) is the set of solutions to (12).

We have shown that equations of the form (9), involving sums of the first \( n \) powers of the \( n \) unknowns, can be solved by solving the \( n \)-order polynomial
equation (8) where the $p_i$'s therein are derived from (6) using (9). Since the resulting polynomial is of order $n$, the same as the number of unknowns, there are no solutions to (8) other than those given by the polynomial; and since the polynomial always has $n$ solutions equations (8) are always consistent. Because computer programs for calculating determinants and for solving polynomial equations are more readily available than are programs for solving equations (9) directly, this procedure for obtaining a solution seems more feasible computationally, especially whenever $n$ is at all large. It is also useful because sets of equations that are variations of (9) can be adapted to the same method of solution. This is now discussed.

Variations on equations (9)

Two variations of interest are when there are more equations than unknowns, and when the equations do not involve a sequence of powers. We consider these variations both separately and in combination.

Suppose the equations to be solved are, similar to (9),

$$s_i = r_i \quad \text{for } i = 1, 2, \ldots, n, \ldots, k$$

where $k$ exceeds $n$, the number of $x$'s involved in the $s_i$. There are now more equations than unknowns, and they have a solution only if they are consistent. Equation (7) readily provides the necessary test. Since there are only $n$ variables all values of $p_t$ for $t > n$, must, as in (7), be zero. Thus the equations are consistent if, for every value of $t > n$, the value of $p_t$ calculated from (6) using $r_i$ for $s_i$, is zero.

Examples. Consider

$$x_1 + x_2 = 7$$
$$x_1^2 + x_2^2 = 25$$
$$x_1^3 + x_2^3 = 91$$

Since $n = 2$, $p_3$ must be zero, and from (4) and (13)

$$6p_3 = 343 - 525 + 182 = 0$$
and hence the equations are consistent. Their solution can be obtained from the first 2 \((=n)\) of them. By (3) and (8) we have

\[ y^2 - 7y + 12 = 0 \]

i.e. equations (13) are consistent and have the solution \((3, 4)\). But for the equations

\[
\begin{align*}
x_1 + x_2 &= 7 \\
x_1^2 + x_2^2 &= 29 \\
x_1^3 + x_2^3 &= 132 ,
\end{align*}
\]

\[ 6p_3 = 343 - 609 + 264 = 2 \neq 0 , \]

and so they are not consistent.

The test must be applied in turn to each value of \(p_t\), for \(t > n\). Thus for the equations

\[
\begin{align*}
x_1 + x_2 &= 3 \\
x_1^2 + x_2^2 &= 5 \\
x_1^3 + x_2^3 &= 9 \\
x_1^4 + x_2^4 &= 16 ,
\end{align*}
\]

we have from (4)

\[ 6p_3 = 27 - 45 + 18 = 0 \]

and from (5)

\[ 24p_4 = 81 + 216 - 270 + 75 - 96 = 6 . \]

Hence \(p_3 = 0\) but \(p_4 \neq 0\). Thus the first three equations of (14) are consistent but all four of them are not.

We now consider equations similar to (9) only in which the sums of powers of the \(n\) unknowns are not sums of the first \(n\) powers but sums of any \(n\) powers, i.e.
where $k_1, \ldots, k_n$ are $n$ positive integers with $k_{j+1} > k_j$ for all $j$. Now for $t > n$, $p_t = 0$, and the equations

$$p_t = 0 \quad \text{for } t = n+1, n+2, \ldots, k_n$$

involve $k_n$ values $s_i, i = 1, 2, \ldots, k_n$, and $n$ of these are given in (15). Therefore (16) can be solved for the $(k_n - n)$ unknown $s$-values. These, together with (15), provide numerical values of $s_i$ for $i = 1, 2, \ldots, n$ and so now equations (9) can be solved as before, using (6) and (8).

**Examples.** From

$$x_1 + x_2 = 5$$
$$x_1^2 + x_2^2 = 35$$

(16) is $p_3 = 0$, which from (4) is

$$s_1^3 - 3s_1s_2 + 2s_3 = 0 ;$$

and from (17) this is

$$125 - 15s_2 + 70 = 0$$

giving $s_2 = 195/15 = 13$.

Therefore (17) is solved from

$$x_1 + x_2 = 5$$
$$x_1^2 + x_2^2 = 13$$

which, from (3) and (8), has solutions given by

$$y^3 - 5y + 6 = 0 ,$$

namely 2 and 3.
The above example leads to but a single solution of the given equations. This is not always the case. Suppose that instead of (17) solutions are sought to

\[ x_1^2 + x_2^2 = 13 \]
\[ x_1^3 + x_2^3 = 35. \]  

Equation (16) is still \( p_3 = 0 \) but from (4) and (19) this is now

\[ s_1^2 - 39s_1 + 70 = 0 \]

i.e. \((s_1 - 5)(s_1 - 14) = 0\),

which has three solutions for \( s_1 \), namely 5, 2 and -7. Accordingly, solutions to (19) can be found by solving

\[ x_1 + x_2 = 5 \]
\[ x_1^2 + x_2^2 = 13 \]

which is (18), with solution 2 and 3; or by solving either

\[ x_1 + x_2 = 2 \quad \text{or} \quad x_1 + x_2 = -7 \]
\[ x_1^2 + x_2^2 = 13 \quad x_1^2 + x_2^2 = 13. \]

The polynomial for the first set of equations in (20) is

\[ y^2 - 2y + \frac{1}{2}(4 - 13) = 0 \]
\[ 2y^2 - 4y - 9 = 0 \],

with solutions \( \frac{1}{2}(2 \pm \sqrt{22}) \); and solving the second set in (20) leads to

\[ y^2 + 7y + \frac{1}{2}(49 - 13) = 0 \]

having solutions \( \frac{1}{2}(-7 \pm i\sqrt{23}) \). Hence equations (19) have three sets of solutions, namely 2 and 3, \( \frac{1}{2}(2 \pm \sqrt{22}) \) and \( \frac{1}{2}(-7 \pm i\sqrt{23}) \), with the respective values of \( s_1 = x_1 + x_2 \) being 5, 2 and -7. In all three cases \( s_2 = x_1^2 + x_2^2 = 13 \) and \( s_3 = x_1^3 + x_2^3 = 35 \).
Another example of multiple solutions is the following. Consider solving

\[ x_1^3 + x_2^3 = 0 \]
\[ x_1^4 + x_2^4 = 2 \]  

Equations (16) are

\[ p_3 = 0 , \]

leading to \( s_1^3 = 2s_1s_2 = 0 \), from (4) and (21) and

\[ p_4 = 0 , \]

leading to \( s_1^4 - 6s_1^2s_2 + s_2^2 - 12 = 0 \) from (5) and (21). From the first of these either \( s_1 = 0 \) or \( s_1^3 = 2s_2 \). The latter implies \( (x_1 - x_2)^2 = 0 \), i.e., \( x_1 = x_2 \), which is clearly inconsistent with (21). We therefore use \( s_1 = 0 \) and using this with (7) and (21) in (5) get

\[ 3s_2^2 = 12 , \]

giving \( s_2 = \pm 2 \). Hence solutions to (21) come from solving either

\[ x_1 + x_2 = 0 \quad \text{or} \quad x_1 + x_2 = 0 \]
\[ x_1^2 + x_2^2 = 2 \quad x_1^2 + x_2^2 = -2 \]

The first of these has the solution \((1, -1)\) and the second is satisfied by \((i, -i)\). Thus (21) has two solutions, \((1, 1)\) and \((i, -i)\).

Finally we consider the situation of more equations than unknowns, involving sums of \( q \) powers of the unknowns, \( q > n \). Similar to (15) they can be specified as

\[ s_i = x_i \quad \text{for} \quad i = k_1, k_2, \ldots, k_q \quad \text{with} \quad q > n \]  

(22)

where \( k_1, k_2, \ldots, k_q \) are positive integers with \( k_{j+1} > k_j \) for all \( j \). We now have

\[ p_t = 0 \quad \text{for} \quad t = n+1, n+2, \ldots, k_q \]  

(23)

involving \( k_q \) \( s \)-values of which \( q \) are given by (22). Thus (23) represents \( k_q - n \) equations in \( k_q - q \) \( s \)-values, with \( k_q - n \) being greater than \( k_q - q \).
The consistency of equations (22) is tested by solving any \( k - q \) of the \( k - n \) equations in (23) and seeing that the solutions satisfy the remaining \( q - n \) equations. If they do, equations (22) are consistent, numeric values for \( s_i \) for \( i = 1, 2, ..., n \) will be available and equations (9) can be solved as before. Their solutions will satisfy (22).

**Example.** Suppose the given equations are

\[
\begin{align*}
  x_1 + x_2 &= 3 \\
  x_1^2 + x_2^2 &= 5 \\
  x_1^4 + x_2^4 &= 17
\end{align*}
\]

We have, for (23),

\[
\begin{align*}
  p_3 &= 0, \text{ implying } s_1^3 - 3s_1s_2 + 2s_3 = 0 \\
  p_4 &= 0, \text{ implying } s_1^4 + 8s_1s_3 - 6s_1s_2 + 3s_2^2 - 6s_4 = 0,
\end{align*}
\]

and from (24) these simplify to

\[
2s_3 = 18
\]

and

\[
2s_3 = 216.
\]

Both of these are satisfied by \( s_3 = 9 \), hence (24) are consistent and can be solved by solving

\[
\begin{align*}
  x_1 + x_2 &= 3 \\
  x_1^2 + x_2^2 &= 5.
\end{align*}
\]

The solution satisfies

\[
y^2 - 3y + 2 = 0
\]

and is clearly 1 and 2.

**Reference**