Improvements to the Kato Method for Finding Square Roots

Revision and Rounding

Professor Fukutaro Kato's method for finding square roots is ideally suited for use on the abacus. The method is clearly described here: http://webhome.idirect.com/~totton/soroban/katoSq/

However, as with most square root methods, it is sometimes necessary to revise a root digit if our first estimate was too large or small. This can be particularly confusing for revising downward, because if our estimate was too large, at some point during the ensuing subtractions from the remainder, the remainder will become negative and it can be difficult to remember how much must be added back to each rod to revise. I will show a straightforward way to revise both upward and downward. Another difficulty with every square root method that I'm aware of is that we are left with a remainder, but with no obvious way to know whether the root developed to that point needs to be rounded up except to do the calculations required to find one more root digit and then round up if that root digit is greater than or equal to five. I will describe an easier method to determine if rounding up is needed.

Revision upward

Just after estimating a new root digit (except for the very first one) and performing the subtractions of the root digit products and one-half the square of the new root digit from the remainder, it may not be obvious if we have underestimated the root digit,especially if the remainder is only a little larger than the root. In the worst case, to check for undershoot after completing the subtractions, we might have to subtract all the root digits except for the last one from the remainder, add back half the square of the underestimated root digit, and then see if there is enough remainder left over to subtract half the square of the corrected root digit.

There is an easier way to check whether our estimate was too low. First, we complete all the subtractions pertaining to a root digit all the way through the subtraction of one half the square of that digit. Then, if we have developed 'n' root digits including the one for which we just completed the subtractions, we simply compare the remainder, beginning at the first rod to the right of the newly developed root digit and comprising n+2 digits including leading zeros, to the n+1 digit number composed of the root with five appended to the right end placed with the most significant digit of the root at the second rod to the right of the newly developed root digit. If the remainder is greater than or equal to the root with five appended, we subtract the root with five appended from the remainder and revise the root digit up by one.

Here's a very simple example: $\sqrt{625} = 25$

006250   set the problem
202250   find the first digit of the root (2) and subtract its square from 06
201125   halve the remainder
240325   (under)estimate the next root digit as 4, subtract 2 x 4 = 8 from 11
240245   subtract $4^2/2$ = 08 from 32

Now we have completed the subtractions for the second root digit. We see that the four digit remainder 0245 is greater than or equal (equal in this case) to the three digit number composed of the root with 5 appended which for the purpose of the comparison is placed with the most significant digit at the second rod to the right of the root, or in this case at the location of the two in the remainder 0245. Hence we can subtract the root with 5 appended and then increment the last root digit to 25.

250000   subtract 245 from 0245 and then increment the last root digit

Note that the Kato remainder is always half of the result of subtracting the square of the root from the radicand. In this case, 0245 or 24.5, is half of the remainder after subtracting the square of the root 24^2 = 576 from the radicand 625.

Here's a more complicated example: $\sqrt{45}$ = 6.70820.....

0450000000000   set the problem
6090000000000   find the first root digit, 6, and subtract its square, 36 from 45
6045000000000   halve the remainder

6703000000000   45/6 = 7+, estimate the next root digit as 7, subtract 7x6 = 42 from 45
6700550000000   subtract $7^2/2$ = 24.5 from 30.0

The remainder, 0055, is less than 675 (the root with 5 appended) so 7 is the correct root digit

6700550000000   05/6 is less than 1, so 0 will be the next root digit
The remainder, 05500, is less than 6705 (the root with 5 appended), so 0 is correct

6708550000000   55/6 = 9+, be cautious and estimate 8, not 9
6708070000000   subtract 8x6 = 48 from 55
6708014000000   subtract 8x7 = 56 from 70
6708014000000   subtract 8x0 = 00 from 40
6708013680000   subtract $8^2/2$ = 32 from 1400

The remainder, 013680, is less than 67085 (the root with 5 appended), so 8 is correct

6708113680000   13/6 = 2+, be cautious and estimate 1, not 2
6708107680000   subtract 1x6 = 06 from 13
6708106980000   subtract 1x7 = 07 from 76
6708106980000   subtract 1x0 = 00 from 98
6708106972000   subtract 1x8 = 08 from 80
6708106971950   subtract $1^2/2$ = 00.5 from 20.0

The remainder 0697195 is greater than 670815 (the root with 5 appended), so subtract 670815 from 697195 and increment the root digit from 1 to 2

6708200263800   now the remainder 0026380 is less than 670825 so go on to the next root digit

If we had chosen 2 originally as the root digit, we would have had:

6708213680000   13/6 = 2+, estimate 2 for the root digit
6708201680000   subtract 2x6 = 12 from 13
6708200280000   subtract 2x7 = 14 from 16
6708200280000   subtract 2x0 = 00 from 28
6708200264000   subtract 2x8 = 16 from 80
6708200263800   subtract $2^2/2$ = 02 from 40

which is exactly where we ended up after revising, so we have indeed revised correctly.

Why does this work? Assume we have a number S, the square root of which is to be found, r is the root developed up to some point, and x is half of the remainder after subtracting the square of that developed root from S (remember that the Kato method works with half the "true" remainder). Now, for the sake of simplicity in the following, we assume that if necessary we have multiplied S by some even power of ten to make r an integer. This does not affect the results as the decimal point can be restored to its original location later.
So we have:

$S = r^2+2x$, or
$S-r^2 = 2x$, or
$x = 1/2(S-r^2)$

It is apparent that if we have underestimated the root by one and subtracted $(r-1)^2$ rather than r^2 from S, our half-remainder would be:
$1/2[S-(r-1)^2] = 1/2[r^2+2x-(r-1)^2] = 1/2[r^2+2x-r^2+2r-1] = x+r-1/2$
that is, our half-remainder x would be too large by r-1/2.

If we then revise by subtracting from this (r-1)+1/2 = (r-1)+0.5 (the underestimated root with 5 appended), we are left with a half-remainder of x, which is exactly what we would have had if we had originally estimated the root correctly and subtracted r^2 rather than $(r-1)^2$.

Revision downward

Revising down can be even more difficult than revising up because if we have overestimated the root digit the remainder will become negative at some point during our subtractions and we may not be easily able to recall how much we need to add back in order to return to the starting point. Instead of trying to reverse direction and begin revising before we have completed all the subtractions it seems easier to complete the subtractions, allowing the remainder to become negative, then revise the root digit downward by one and add the corrected root with five appended to the negative remainder, similarly to the way we subtracted the uncorrected root with five appended when we underestimated the root digit.

Here's an example: $\sqrt{3364} = 58$

abcdef
033640 set the problem
533640 the first root digit is 5
508640 subtract the square of the first digit from 33
504320 halve the remainder
594320 overshoot the next digit, 9 instead of 8
599820 subtract 45 from 43 (98 is the 10's complement of 02)
599415 subtract $9^2/2 = 40.5$ from 82
589415 decrement the root digit to correct the overshoot
580000 add 585 to 415 (start at the second rod from the root)

Notice that when we subtracted 45 from 43 on rods 'cd' we first subtracted four from four on rod 'c', leaving a remainder of 03 on rods 'cd', then when we subtracted five from three on rod 'd' we needed to borrow from the zero on rod 'c', requiring a borrow in turn from rod 'b' which we can't do because rod 'b' is part of the root, not part of the remainder. So we just assumed that we borrowed one from “somewhere” and then returned it when we added 585 to 415 causing a carry into rod 'c', making it a zero and causing a carry into “somewhere” (because we can't carry into rod 'b' which is part of the root) thus replacing the earlier borrow. Everything works out in the end as long as we remember not to borrow from or carry into the root.

Here's another example: $\sqrt{121801} = 349$

01218010	set the problem
31218010	the first root digit = 3
30318010	subtract the square of first root digit
30159005	halve the remainder
35159005	overshoot the second root digit, 5 instead of 4
35009005	subtract 15 from 15
35996505	subtract 125 from 090 - the remainder goes negative - continue borrows leftward up to the most significant digit of the remainder, but not into the root
34996505	correct the overshoot by decrementing the last root digit
34031005	and add 345 to 965 - continue carries leftward up to the most significant digit of the remainder, but not into the root
34931005	the next root digit is 9
34904005	subtract 27 from 31
34900405	subtract 36 from 40
34900000	subtract 405 from 405

finished: the root is 349

Why does it work? As in the explanation of why the undershoot method works, we again assume that if necessary we have multiplied S by an even power of ten so that r is an integer. As before, if $S = r^2+2x$ and we overestimate the root digit and subtract $(r+1)^2$ instead of r^2, our half-remainder would be:
$1/2[S-(r+1)^2] = 1/2[r^2+2x - (r+1)^2] = 1/2[r^2+2x - r^2-2r-1] = x-r-1/2$.
If we then revise by adding to this $r+1/2 = r+0.5$ (the corrected root with 5 appended), we are left with a half-remainder of x, which is exactly what we would have had if we had originally estimated the root correctly and subtracted r^2 rather than $(r+1)^2$.

Rather than trying to remember that when we revise an undershoot we correct the remainder first, then correct the root, but when we revise an overshoot we correct the root first, then correct the remainder; it may be easier to remember this rule: always append five to the smaller of the uncorrected or the corrected root. This will be the underestimated root for an undershoot and the corrected root for an overshoot. Following this rule, it doesn't matter whether the root or the remainder is corrected first.

Rounding

The instructions for any square root method will leave us with a remainder if the radicand is not a perfect square. Usually we don't really care about the remainder, but what we really want to know is the root to a certain number of significant digits, say 'n' digits. If we find 'n' digits and the 'n+1'st digit is equal to or greater than five then we should increment the 'n' digit root. Of course we can use the obvious method of calculating 'n+1' digits of the root and then rounding up the 'n'th digit or not. But there is a much easier way; just compare the 'n+3' digit “true” remainder (double the “Kato” remainder) beginning at the rod to the right of the 'n'th root digit to the 'n' digit root with 25 appended to make an 'n+2' digit number. If the true remainder is equal to or greater than the 'n' digit root with 25 appended, then increment the root. We can make the comparison mentally just by looking at the remainder and the root or we can simply begin subtracting the root (with 25 appended) from the remainder and see if the result becomes negative, in which case no rounding up is required. If it doesn't become negative then we must round up.

Example: $\sqrt{228520}$ = 478, remainder 36.

If we want 3 significant digits, should we round up to 479 ?

When we finish with the three digit root, our abacus looks like this:

478001800

and after doubling the remainder, we have
478003600

We can see at a glance that the remainder 003600 is less than 47825, so rounding up the root to 479 is not needed.

Alternatively, we can just begin subtracting the root with 25 appended, starting at the second digit to the right of the last root digit:

```
478003600
-   47825
---------
```

This result becomes negative immediately, so we know that rounding up is not needed.

Another example: Find $\sqrt{976}$ = 31.24090.... to six significant digits:

After developing the root to 6 significant digits, the abacus looks like this:

3124090308359500

after doubling the remainder, we have
3124090616719000

To determine if we need need to round up to 31.2410, compare the remainder 061671900 to the root with 25 appended 31240925 or alternatively just subtract the root with '25' appended, starting at the second rod to the right of the last root digit and verify that the result is not negative:

312409061671900
-      31240925

Clearly the remainder is larger than the root with 25 appended, so in this case, rounding up to 31.2410 is required.

Here's why this works:

Assume we have a number S and we want to find the square root of S to 'n' significant figures. After we have found the root 'r' to 'n' digits, whether the next digit is greater than or equal to five will determine if we need to round up the 'n'th digit. (As in the explanations above, we assume that if necessary we had multiplied S by some even power of ten such that r is an integer.) Now consider the case where S is such that the root is right on the dividing line between rounding up the 'n'th digit or not; that is, the 'n' digit root is 'r' and the next, or 'n+1'st digit is exactly five.   Then the 'n' digit root is r+0.5 and we have:

$S = (r+0.5)^2 = r^2+r+0.25$ or
$S-r^2 = r+0.25$

After finding r (the square root of S to 'n' digits) and doubling the half-remainder to get the true remainder $S-r^2$, we are left with a remainder of r+0.25. In the general case where the root is not right on the dividing line of being rounded up, if the remainder is less than r+0.25, the root must have been less than r+0.5 so we don't need to round up. On the other hand, if the remainder is greater than or equal to r+0.25, the root must have been greater than or equal to r+0.5, so we do need to round up the 'n'th digit.

Steve Treadwell
May, 2015