MATH 181 NOTES: FEBRUARY 21, 2024

JESSE LEO KASS

We are going to going to continue examining the unfamiliar material in nineteenth century college algebra textbooks that were commonly used in America. Both the textbook by Davies and the textbook by Ficklin include some unfamiliar advanced material on the theory of equation Sturm's theorem and Horner's method. We just examined Sturm's theorem, and here we will take a look at Horner's method.

As presented in the textbooks, Horner's method is a method for computing the digits of a solution to a polynomial equation $f(x) = 0$. As presented by Ficklin, Horner's method is described in Figure 1. Let's see how this works in the simple example of $f(x) = x^2 - 2$. The first step is to find the integral part of the root. The root question is $\sqrt{2}$ which has

RULE.

I. *Find the integral part of the root by Sturm's Theorem or otherwise.*

II. *Find an equation whose roots shall be less than those of the given equation by the integral part of the required root.*

III. *Divide the independent term of the transformed equation by the coefficient of the adjacent term, change the sign of the quotient and write the first figure of the result as the first figure of the fractional part of the root.*

IV. *Find an equation whose roots shall be less than those of the second equation by the first figure in the fractional part of the required root.*

V. *Divide the independent term of this transformed equation by the coefficient of the adjacent term, change the sign of the quotient, and write the first figure of the result as the second figure of the fractional part of the required root.*

VI. *Continue this process until the root is obtained to the required degree of accuracy.*

FIGURE 1. Ficklin's description of Horner's rule

$r_0 = 1$. The next step is transform the equation to a new equation with roots equal to $\pm\sqrt{2} - 1$, which is one less than the roots of $f(x)$.

Let's examine what happens when the roots of a polynomial are transformed. Suppose that $f(x) = x^2 + ax + b$ is a quadratic polynomial with roots ρ_1, ρ_2, so $f(x) = (x - \rho_1)(x - \rho_2)$. The new polynomial is

$$\begin{aligned} g(x) :=& (x - (\rho_1 - 1))(x - (\rho_2 - 1)) \\ =& x^2 - (\rho_1 + \rho_2 - 2)x + \rho_1\rho_1 - \rho_1 - \rho_2 + 1. \end{aligned}$$

Comparing with the coefficients of

$$\begin{aligned} x^2 + ax + b =& f(x) \\ =& (x - \rho_1)(x - \rho_1) \\ =& x^2 - (\rho_1 + \rho_2)x + \rho_1\rho_2, \end{aligned}$$

we get that $a = -(\rho_1 + \rho_2)$ and $b = \rho_1\rho_2$. We conclude that

$$g(x) = x^2 + (a + 2)x + b + a + 1.$$

In fact, we could have saved some work. Rearranging terms, we have

$$\begin{aligned} g(x) :=& (x - (\rho_1 - 1))(x - (\rho_2 - 1)) \\ =& ((x + 1) - \rho_1)((x + 1) - \rho_2) \\ =& f(x + 1). \end{aligned}$$

Thus we could have just evaluated $f(x + 1)$ instead of working with the roots ρ_1, ρ_2. Applying these formulas to $f(x) = x^2 - 2$, we get

$$g(x) = x^2 + 2x - 1$$

What do we do with this last equation? Rule III is confusingly worded. The "independent" term of the constant coefficient, so Ficklin is saying that the "first figure" of the result is 1/2. Elsewhere in the text, Ficklin explains where this comes from. The idea is to estimate a solution to $x^2 + 2x - 1 = 0$ by a solution to the equation $2x - 1$ obtained by dropping all the nonlinear terms. This last equation is easy to solve because it is linear. We then add the "first figure" to our last estimate to get a better estimate. This yields 1.5. This is indeed a better estimate for $\sqrt{2}$. We have $1^2 = 1$, but $(1.5)^2 = 2.25$.

We can repeat this computation to improve the estimate $\sqrt{2} \approx 2.25$ by computing the polynomial $g_1(x) = g(x + .5)$ and then estimating a root of $g_1(x)$ by looking at the linear terms. In general, we can compute $g(x + \sigma)$ by the formula

$$\begin{aligned} f(x + \sigma) =& (x + \sigma)^2 + a(x + \sigma) + b \\ =& x^2 + (2\sigma + a)x + (\sigma^2 + a\sigma + b). \end{aligned}$$

Figure 2 shows an implementation of the method in an Excel spreadsheet. After about seven steps, the spreadsheet produces garbage because of rounding errors, but all the displayed digits of the estimate of $\sqrt{2}$ are accurate after five steps.

Here's a modern summary of what Ficklin calls Horner's method:

A	B	C	D
Estimate	N-th figure	Coef of x	Constant term
1	1	0	-2
1.5	0.5	2	-1
1.41666667	-0.0833333	3	0.25
1.41421569	-0.002451	2.83333333	0.00694444
1.41421356	-2.124E-06	2.828431373	6.0073E-06
1.41421356	-1.595E-12	2.82842712	4.511E-12
1.41421356	-8.993E-25	2.82842712	2.5437E-24

FIGURE 2. Excel implementation of Horner's rule for $x^2 - 2$

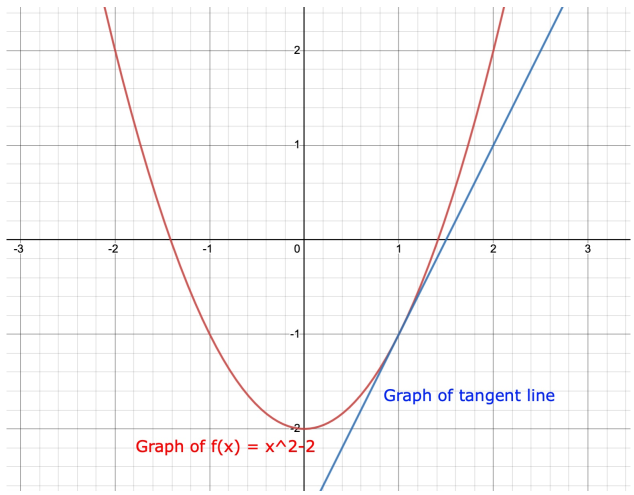

FIGURE 3. Graphical illustration of Newton's method

(1) Guess the integer part r_0 of a solution $f(x) = 0$. Set $n = 0$, $f_0(x) = f(x)$, and $y_n = r_0$.
(2) Compute the polynomial $f_{n+1}(x) := f_n(x + y_n)$.
(3) Let $y_n = -a_n/b_n$ where a_n is the constant term of f_{n+1} and b_n the coefficient of x. The new root estimate is $r_{n+1} = r_n + y_n$.
(4) Return to step 2 with n replaced by $n + 1$

Today, Horner's method is rarely taught at the undergraduate level. More common is teaching Newton's method. Newton's method is best explained in geometric terms. Starting with an initial guess r_0, we replace the graph of $f(x)$ with its tangent line at $(r_0, f(r_0))$. Then we calculate where the tangent line crosses the x-axis and use this as our improved estimate r_1 for a root. To improve the estimate, repeat. Figure 3 shows the basic graphical idea.

Recall from calculus that the tangent line to $y = f(x)$ at $(r_0, f(r_0))$ is the line that passes through $(r_0, f(r_0))$ and has slope $f'(r_0)$. The equation for this line is

$$y = f'(r_0)x + f(r_0) - r_0 f'(r_0).$$

Setting $y = 0$ and solving for x, we get $x = r_0 - f(r_0)/f'(r_0)$. With the formula, we can forget about the geometry. This yields the following rule:

A	B	C
Estimate	f(x)	Derivative of f(x)
1	-1	2
1.5	0.25	3
1.41666667	0.00694444	2.833333333
1.41421569	6.0073E-06	2.828431373
1.41421356	4.5106E-12	2.828427125
1.41421356	0	2.828427125
1.41421356	0	2.828427125

FIGURE 4. Excel implementation of Newton's method for $x^2 - 2$

(1) Guess the integer part r_0 of a solution $f(x) = 0$. Set $n = 0$, $f_0(x) = f(x)$, and $y_n = r_0$.
(2) Set $r_{n+1} = r_n - f(r_0)/f'(r_0)$.
(3) Return to step 2 with n replaced by $n + 1$

An Excel implementation for $f(x) = x^2 - 2$ and $r_0 = 1$ is displayed in Figure 4.

Notice anything? Both Newton's method and Horner's method produce the same estimates! A closely look at the two methods explains what is going on. If

$$f(x) = x^2 + ax + b,$$

then the next polynomial used in Horner's method is

$$\begin{aligned} g(x) =& f(x + r_0) \\ =& x^2 + (2r_0 + a)x + (r_0^2 + ar_0 + b) \\ =& x^2 + f'(r_0)x + f(r_0). \end{aligned}$$

Thus $y = -f(r_0)/f'(r_0)$ and the next estimate produced by Horner's method if $r_0 + y = r_0 - f(r_0)/f'(r_0)$, exactly the estimate used in Newton's method.

Our close look at $f(x) = x^2 - 2$ suggests that the two methods only coincided because we worked with a quadratic polynomial. Let's look at what happens if we look at a polynomial of higher degree. When we studied Sturm's theorem, we looked at the polynomial $f(x) = 2x^4 - 13x^2 + 10x - 19$. As an application of Sturm's theorem, we found that $f(x)$ has a unique positive root that is close to $r_0 = 2$.

To automate the computation, let's see how a degree 4 polynomial transforms when we transform the roots. The relevant computation is the following one: if

$$f(x) = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0),$$

then

$$\begin{aligned} f(x + r_0) =& x^4 + (4r_0 + a_3)x^3 + (6r_0^2 + 3a_3r_0 + a_2)x^2 \\ &+ (4r_0^3 + 3a_3r_0^2 + 2a_2r_0 + a_1)x + (r_0^4 + a_3r_0^3 + a_2r_0^2 + a_1r_0 + a_0). \end{aligned}$$

The resulting estimates are displayed in Figures 5 and 6.

	A	B	C	D
	Estimate	f(x)	Derivative of f(x)	
	2	-19	126	
	2.15079365	-14.830789	145.515715	
	2.25271246	-10.938717	160.0254854	
	2.32106856	-7.7777054	170.3832234	
	2.36671686	-5.40015	177.5890896	
	2.39712498	-3.6916398	182.5202824	
	2.41735089	-2.4981212	185.8590933	

FIGURE 5. Excel implementation of Newton's method for $2x^4 - 13x^2 + 10x - 19$

	A	B	C	D	E	F
1	Estimate	N-th figure	Coefficient of x^3	Coefficient of x^2	Coef of x	Constant term
2	2	2	0	-6.50	5	-9.5
3	2.86363636	0.86363636	8	17.5	11	-9.5
4	2.55957099	-0.3040654	11.45454545	42.70247934	61.7047333	18.7622729
5	2.4658969	-0.0936741	10.23828395	32.80842186	38.8007079	3.63462096
6	2.45740519	-0.0084917	9.863587598	29.98388511	32.9203398	0.27954987
7	2.45733867	-6.652E-05	9.829620771	29.73304169	32.4132424	0.00215608
8	2.45733867	-4.059E-09	9.829354697	29.73108016	32.4092869	1.3156E-07
9						

FIGURE 6. Excel implementation of Horner's rule for $2x^4 - 13x^2 + 10x - 19$

Now we are seeing a big difference between the two methods. Moreover, Horner's method has produced a better estimate. The last Horner's estimate is correct to at least four digits, while Newton's only has the first digit correct.

One final example. Figure 7 displays one of the examples from Davies's textbook, and Figure 8 shows an implementation of Horner's method in Excel. If you look closely, you'll see that Excel produces a slightly different answer than Davies, presumably because Excel introduces rounding error.

Why did the textbooks by Davies and Ficklin present Horner's method rather than Newton's method. From the Excel tables, we see that Horner's method appears to provide better estimates. For the polynomial $f(x) = 2x^4 - 13x^2 + 10x - 19$, applying Horner's method three times produces a better estimate we get by applying Newton's method five times. We should be careful here because each step of Horner's method involved more computations. To apply Newton's method, we only compute $f(r_n)$ and $f'(r_n)$; Horner's requires us to compute those two quantities as well as the $\deg(f) - 2$ additional ones. A

2. Find the roots of the equation $x^4 - 12x^2 + 12x - 3 = 0$.

$$\textit{Ans.} \begin{cases} +\ 2.858083308163 \\ +\ .606018306917 \\ +\ .443276939605 \\ -\ 3.907378554685. \end{cases}$$

FIGURE 7. Example from the Davies textbook

	A	B	C	D	E	F
1	Estimate	N-th figure	Coefficient of x^3	Coefficient of x^2	Coef of x	Constant term
2	3.0000000000000	3	0	-12	12	-3
3	2.8750000000000	-0.125	12	42	48	6
4	2.8583645555327	-0.016635444	11.5	37.59375	38.0546875	0.633056641
5	2.8580833877196	-0.000281168	11.43345822	37.02148759	36.81343907	0.010350754
6	2.8580833081794	-7.95402E-08	11.43233355	37.01184391	36.79262328	2.92649E-06
7						

FIGURE 8. Excel implementation of the example

textbook more advanced than those by Davies and Ficklin would spend more time examining how to rapidly compute all the $\deg(f)$ quantities that used in each step of Horner's method.

A major disadvantage of Horner's method is that it only applies to polynomials, while Newton's method applies quite generally to differentiable functions. In the context of the books by Davies and Ficklin, this is not a serious problem. Those books are *algebra* textbooks, so the authors don't even consider functions more complicated than polynomials. Moreover, while Newton's method can be described without using any geometry, it seems completely unmotivated unless one introduces the geometric meaning of the derivative.

There are still a number of interesting questions that have gone unanswered. Here's a few I thought of:

(1) What did Horner have to do with Horner's rule?
(2) What did Newton have to do with Newton's method?
(3) How to prove that Horner's theorem works? Why do Ficklin and Davies omit a proof?
(4) Did American college students actually use Horner's method after graduation? If so, what did they use it for?
(5) Did American college students learn Newton's method in calculus?
(6) When/why did Horner's rule get removed from the college curriculum?

You can probably come up with some other interesting questions yourself. We'll try to answer some of these questions next class.

DEPT. OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT SANTA CRUZ