Newton's Method

Find root of $f(x) = 0$ through successive approximation e.g., $f(x) = x^2 - a$
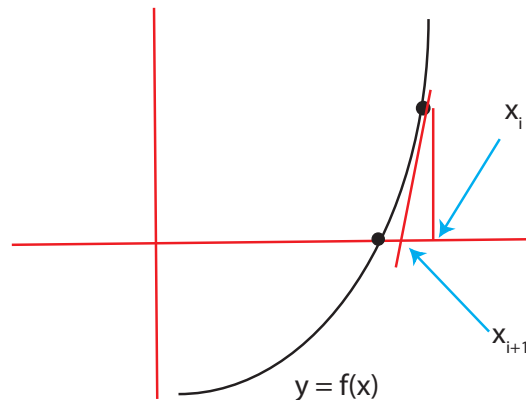

Figure 1: Newton's Method.

Tangent at $(x_i, f(x_i))$ is line $y = f(x_i) + \underline{f'(x_i)} \cdot (x - x_i)$ where $f'(x_i)$ is the derivative. $x_{i+1} =$ intercept on x-axis

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

Square Roots

We want to find the square root of a number a through successive approximation using Newton's Method, so that our answer x will converge on $a^{1/2}$. It looks like $f(x) = a^{1/2} - x$ has a root at $x = a^{1/2}$. Let's try to use it with Newton's Method.

$$
\begin{aligned}
f(x) &= a^{1/2} - x \\
f'(x) &= -1 \\
\chi_{i+1} &= \chi_i - \frac{f(\chi_i)}{f'(\chi_i)} \\
&= \chi_i - \frac{a^{1/2} - \chi}{-1} \\
&= \chi_i + a^{1/2} - \chi_i \\
&= a^{1/2}
\end{aligned}
$$

The choice of this function was unsuccessful, since we're unable to compute x_{i+1} using x_i. Let's try squaring both sides of $x = a^{1/2}$. This time, it looks like we get something usable:

$$f(x) = x^2 - a$$

$$\chi_{i+1} = \chi_i - \frac{(\chi_i{}^2 - a)}{2\chi_i} = \frac{\chi_i + \dfrac{a}{\chi_i}}{2}$$

Error Analysis of Newton's Method

Suppose $X_n = \sqrt{a} \cdot (1 + \epsilon_n)$ ϵ_n may be + or -
Then,

$$
\begin{aligned}
X_{n+1} &= \frac{X_n + a/X_n}{2} \\
&= \frac{\sqrt{a}(1 + \epsilon_n) + \frac{a}{\sqrt{a}(1+\epsilon_n)}}{2} \\
&= \sqrt{(a)}\frac{\left((1 + \epsilon_n) + \frac{1}{(1+\epsilon_n)}\right)}{2} \\
&= \sqrt{(a)}\left(\frac{2 + 2\epsilon_n + \epsilon_n{}^2}{2(1 + \epsilon_n)}\right) \\
&= \sqrt{(a)}\left(1 + \frac{\epsilon_n{}^2}{2(1 + \epsilon_n)}\right)
\end{aligned}
$$

Therefore,

$$\epsilon_{n+1} = \frac{\epsilon_n{}^2}{2(1 + \epsilon_n)}$$

Quadratic convergence, as $\sharp$ correct digits doubles each step.

Complexity of Computing Square Roots

We apply a first level of Newton's method to solve $f(x) = x^2 - a$. Each iteration of this first level requires a division. If we set the precision to d digits right from the beginning, then convergence at the first level will require $\lg d$ iterations. This means the complexity of computing a square root will be $\Theta(d^\alpha \lg d)$ if the complexity of multiplication is $\Theta(d^\alpha)$, given that we have shown that the complexity of division is the same as the complexity of multiplication.

However, we can do better, if we recognize that the number of digits of precision we need at beginning of the first level of Newton's method starts out small and then grows. If the complexity of a d-digit division is $\Theta(d^\alpha)$, then a similar summation to the one above tells us that the complexity of computing square roots is $\Theta(d^\alpha)$.

Termination

Iteration: $\chi_{i+1} = \lfloor \dfrac{\chi_i + \lfloor a/\chi_i \rfloor}{2} \rfloor$

Do floors hurt? Does program terminate? (α and β are the fractional parts below.)

Iteration is

$$
\begin{aligned}
\chi_{i+1} &= \frac{\chi_i + \frac{a}{\chi_i} - \alpha}{2} - \beta \\
&= \frac{\chi_i + \frac{a}{\chi_i}}{2} - \gamma \quad \text{where } \gamma = \frac{\alpha}{2} + \beta \text{ and } 0 \le \gamma < 1
\end{aligned}
$$

Since $\dfrac{a+b}{2} \ge \sqrt{ab}$, $\dfrac{\chi_i + \frac{a}{\chi_i}}{2} \ge \sqrt{a}$, so subtracting γ always leaves us $\ge \lfloor \sqrt{a} \rfloor$. This won't stay stuck above if $\epsilon_i < 1$ (good initial guess).

Cube Roots

Now that we've seen square roots, let's calculate cube roots. This time, x will converge on $a^{1/3}$, instead. We can use $f(x) = a - x^3$:

$$
\begin{aligned}
f(x) &= x^3 - a \\
f'(x) &= 3x^2 \\
\chi_{i+1} &= \chi_i - \frac{f(\chi_i)}{f'(\chi_i)} \\
&= \chi_i - \frac{1}{3}\left(\chi_i - \frac{a}{\chi_i^2} \right) \\
&= \frac{2}{3}\chi_i + \frac{1}{3}\frac{a}{\chi_i^2}
\end{aligned}
$$

Suppose $x_{i+1} = a^{1/3} \cdot (1 + \epsilon_i)$ ϵ_i may be + or -
Then,

$$
\begin{aligned}
\chi_{i+1} &= \frac{2}{3}\chi_i + \frac{1}{3}\frac{a}{\chi_i^2} \\
&= \frac{2}{3}a^{1/3}(1 + \epsilon_i) + \frac{1}{3}\frac{a}{(a^{1/3}(1 + \epsilon_n))^2} \\
&= \frac{a^{1/3}}{3}\left(2(1 + \epsilon_i) + \frac{1}{(1 + \epsilon_i)^2}\right) \\
&= \frac{a^{1/3}}{3}\left(\frac{1}{(1 + \epsilon_i)^2}\right)\left(2(1 + 3\epsilon_i + 3\epsilon_i^2 + \epsilon_i^3) + 1\right) \\
&= \frac{a^{1/3}}{3}\frac{3 + 6\epsilon_i + 6\epsilon_i^2 + 2\epsilon_i^3}{1 + 2\epsilon_i + \epsilon_i^2} \\
&= \frac{a^{1/3}}{3}\left(3 + \frac{3\epsilon_i^2 + 2\epsilon_i^3}{1 + 2\epsilon_i + \epsilon_i^2}\right) \\
&= a^{1/3}\left(1 + \frac{\epsilon_i^2 + \frac{2}{3}\epsilon_i^3}{(1 + \epsilon_i)^2}\right)
\end{aligned}
$$

Notice that the cubic term gets dominated by the quadratic term in the numerator, and that the denominator is very close to 1. Therefore, we have quadratic convergence here, just like in the square root case. The running time analysis is very similar to the square root case as well: the only difference is one additional multiplication in the denominator, so the running time stays the same as the running time of a multiplication.

6.006 Introduction to Algorithms

Fall 2011