CS412: Lecture #4

Mridul Aanjaneya

January 29, 2015

Fixed point iteration (cont'd)

In the previous lecture, we described the fixed point iteration

$$x_{k+1} = g(x_k)$$

as a method for solving the nonlinear equation $f(x) = 0$. The *iteration function* $g(x)$ was not uniquely defined, but we stated certain conditions that would need to be satisfied in order for this methodology to be effective. First, we asked that, if the iteration converges, the limit must be a solution to $f(x) = 0$. We saw that this can be guaranteed if

$$f(x) = 0 \Leftrightarrow x = g(x)$$

This condition does not uniquely determine $g(x)$; in fact, we saw several definitions of $g(x)$ that satisfy the above equivalence. Unfortunately, not all of them will lead to an effective method. For example, consider the nonlinear equation $f(x) = x^2 - a = 0$ (solution: $\pm\sqrt{a}$) and the function $g(x) = a/x$. We can easily verify that

$$x = g(x) = \frac{a}{x} \Leftrightarrow x^2 = a \Leftrightarrow x^2 - a = 0 = f(x)$$

However, the iteration $x_{k+1} = g(x_k) = a/x_k$ yields,

$$x_1 = \frac{a}{x_0}, \quad x_2 = \frac{a}{x_1} = \frac{a}{a/x_0} = x_0$$

Thus, the sequence alternates forever between the values $x_0, x_1, x_0, x_1, \ldots$ regardless of the initial value. Other choices of $g(x)$ may also create divergent sequences, often regardless of the value of the initial guess.

Fortunately, there are ways to ensure that the sequence $\{x_k\}$ converges, by making an appropriate choice of $g(x)$. We will use the following definition:

Definition: A function $g(x)$ is called a *contraction in the interval* $[a, b]$, if there exists a number $L \in [0, 1)$ such that

$$|g(x) - g(y)| \leq L|x - y|$$

for any $x, y \in [a, b]$.

Examples:

- $g(x) = x/2$:
$$|g(x) - g(y)| = \frac{1}{2}|x - y|$$
for any $x, y \in \mathbb{R}$.
- $g(x) = x^2$, in $[0.1, 0.2]$:
$$|g(x) - g(y)| = |x^2 - y^2| = |x + y||x - y| \leq 0.3|x - y|$$
for $x, y \in [0.1, 0.2]$ (in this case this condition is essential!)

If we can establish that the function g in the fixed point iteration $x_{k+1} = g(x_k)$ is a contraction, we can show the following:

Theorem: Let a be the actual solution to $f(x) = 0$, and assume $|x_0 - a| < \delta$, where δ is an arbitrary positive number. If g is a contraction on $(a - \delta, a + \delta)$, the fixed point iteration is guaranteed to converge to a.

Proof. Since a is the solution, we have $a = g(a)$. Thus,
$$\begin{aligned} |x_1 - a| &= |g(x_0) - g(a)| \leq L|x_0 - a| < L\delta \\ |x_2 - a| &= |g(x_1) - g(a)| \leq L|x_1 - a| < L^2\delta \\ &\vdots \\ |x_k - a| &< L^k\delta \end{aligned}$$
Since $L < 1$, we have $\lim_{k\to\infty} |x_k - a| = 0$, i.e., $x_k \to a$. $\square$

In some cases, it can be cumbersome to apply the definition directly to show that a given function g is a contraction. However, if we can compute the derivative $g'(x)$ we have a simpler criterion:

Theorem: If g is differentiable and a number $L \in [0, 1)$ exists such that $|g'(x)| \leq L$ for all $x \in [a, b]$, then g is a contraction on $[a, b]$.

Proof. Let $x, y \in [a, b]$ and, without loss of generality, assume $x < y$. The mean value theorem states that
$$\frac{g(x) - g(y)}{x - y} = g'(c) \quad \text{for some } c \in (x, y).$$
Now, if $|g'(x)| \leq L$ for all $x \in [a, b]$, then regardless of the exact value of c we have
$$|g'(c)| \leq L \Rightarrow \left|\frac{g(x) - g(y)}{x - y}\right| \leq L \Rightarrow |g(x) - g(y)| \leq L|x - y|$$
$\square$

Examples:

- Let $g(x) = \sin(\frac{2x}{3})$. Then

$$|g'(x)| = \frac{2}{3}\left|\cos\left(\frac{2x}{3}\right)\right| \leq \frac{2}{3} < 1$$

  Thus, g is a contraction.

- Let us try to apply the derivative criterion to see if the function

$$g(x) = x - \frac{f(x)}{f'(x)}$$

  which defines Newton's method is a contraction:

$$g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}$$

  Now let us assume that

  - $f(a) = 0$, i.e., a is a solution of $f(x) = 0$,
  - $f'(a) \neq 0$, and
  - f'' is *bounded* near a (for example, if f'' is continuous).

  Then

$$\lim_{x\to a} g'(x) = \frac{f(a)f''(a)}{[f'(a)]^2} = 0$$

  This means that there is an interval $(a-\delta, a+\delta)$ where $|g'(x)|$ is *small* (since $\lim_{x\to a} g'(x) = 0$). Specifically, we can find an $L < 1$ such that $|g'(x)| \leq L$ when $|x - a| < \delta$. This means that g is a contraction on $(a - \delta, a + \delta)$, and if the initial guess also falls in that interval, the iteration is guaranteed to converge to the solution a.

Let us revisit Newton's method once again. The equality we showed previously

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

can give us some insights about certain cases, where convergence is more likely, and others where convergence may be at risk:

- If f'' is small, $g'(x)$ will also tend to be small. In the limit case where $f''(x) = 0$, convergence is instantaneous. Of course, this is of limited interest because it would imply that the equation of interest is in fact linear, or $f(x) = ax + b$. However, when $f''(x) \approx 0$, we can expect very rapid convergence.

- If $f'(x)$ is large, convergence will typically occur more easily. Of course, sometimes this fact coincides with f'' being large, in which case the two factors compete or cancel one another.
- Another consequence is that, when $f'(x) \approx 0$ (i.e., the graph of f is mostly "flat"), convergence will be less certain. Compare this with our intuitive graphical explanation of "flat" tangents in Newton's method.