The tangent approximation

March 10

Let $f$ be a function of $x$. Suppose that we can calculate $f(a)$ and $f'(a)$. Then the tangent line to the graph of $f$ at $a$ is the line passing through the point $(a, f(a))$ whose slope is $f'(a)$. It is given by the formula

$$
\ell_a(x) = f'(a)(x-a) + f(a).
$$

It is a theorem that this is the line which best approximates $f$ near $a$. Although we won’t try to say exactly what this means, we will explain how well it does approximate $f$. The key point is that, for $x$ near $a$, the difference between $f(x)$ and $\ell_a(x)$ is small even compared to $|x-a|$.

As a simple example, consider the function $f(x) = x^2$. Then $f'(a) = 2a$ and $\ell_a(x) = 2a(x-a) + a^2 = 2ax - a^2$. The point is, if we already know $a^2$, this is easier to compute than $x^2$ and is supposed to be near to $x^2$ if $x$ is near $a$. How near? We can compute the difference:

$$
|f(x) - \ell_a(x)| = |x^2 - 2ax + a^2| = |x-a|^2
$$

Note that if $|x-a| < 1$, this is small even compared to $|x-a|$.

Here is a precise statement.

**Theorem:** Suppose that $f'(a)$ exists. Then for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$
|f(x) - \ell_a(x)| \leq \epsilon |x-a|
$$

whenever $|x-a| < \delta$.

For example, if $\epsilon$ is chosen to be .01, the error caused by using $\ell_a(x)$ in place of $f(x)$, will be at most 1% of the the difference between $x$ and $a$.

**Proof:** From the definition of derivative:

$$
f'(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x-a} = \lim_{x \to a} \frac{f(x) - f(a)}{x-a}.
$$
By the definition of a limit, we can find $\delta > 0$ such that
\[
\left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| < \epsilon
\]
whenever $0 < |x - a| < \delta$. Now multiply both sides by the positive number $|x - a|$ to see that
\[
|f(x) - f(a) - f'(a)(x - a)| < \epsilon |x - a|
\]
whenever $0 < |x - a| < \delta$. Note that both sides vanish when $x = a$, so if we replace the “less than” sign by a “less than or equal” sign, the statements remains true for all $x$ with $|x - a| < \delta$. Now if we substitute in the definition of $\ell_a(x)$, we see that
\[
|f(x) - \ell_a(x)| \leq \epsilon |x - a|
\]
whenever $|x - a| < \delta$. This proves the theorem.

Let’s work out an example. Let $f(x) = \sqrt{x}$, for $x > 0$. Then if $a > 0$, $f'(a) = 1/2a^{-1/2}$, so $\ell_a(x) = (x - a)/2\sqrt{a} + \sqrt{a}$. Let’s see if, given $\epsilon$ we can find $\delta$ that works in the above argument. In the calculations below, we shall frequently use our old friend: $A^2 - B^2 = (A + B)(A - B)$.

\[
|f(x) - \ell_a(x)| = \left| \sqrt{x} - \sqrt{a} - \frac{(x - a)}{2\sqrt{a}} \right|
\]
\[
= \left| (\sqrt{x} - \sqrt{a}) \left( 1 - \frac{\sqrt{x} + \sqrt{a}}{2\sqrt{a}} \right) \right|
\]
\[
= \frac{x - a}{\sqrt{x} + \sqrt{a}} \left( \frac{\sqrt{a} - \sqrt{x}}{2\sqrt{a}} \right)
\]
\[
= \left| \frac{(x - a)^2}{(\sqrt{x} + \sqrt{a})^2} \right|
\]

This is still pretty messy. We don’t have to be very clever to get something useful and simple however. Since $\sqrt{x}$ is positive, if we omit it from the denominator we will get something bigger. So we conclude:

\[
|f(x) - \ell_a(x)| \leq \left| \frac{(x - a)^2}{2(\sqrt{a})^3} \right|
\]
In our example, \( a = 25 \), and so \( \sqrt{a} = 5 \) and \( \sqrt[3]{a} = 125 \), so we get
\[
|f(x) - \ell_{25}(x)| \leq \left| \frac{(x - 25)^2}{250} \right| = |x - 25| \frac{|x - 25|}{250}
\]

Conclusion: if we take \( \delta := 250\epsilon \), then if \( |x - 25| < \delta \),
\[
|f(x) - \ell_a(x)| \leq \epsilon |x - a|.
\]

Thus this \( \epsilon \) is a bound for the relative error
\[
\rho := \frac{|f(x) - \ell_a(x)|}{|x - a|}
\]
(which makes sense only if \( x \neq a \)).

Let’s look at some values. The estimates we just did predict that the relative error \( \rho \) is bounded by \( .004|x - 25| \). Since it is impossible to actually write down \( f(x) \) exactly, I have written \( \tilde{f}(x) \) to indicate the approximation given by my calculator.

| \( x \) | \( x - a \) | \( \ell_a(x) \) | \( f(x) \) | \( \rho \) | \(.004|x - a|\) |
|---|---|---|---|---|---|
| 25 | 0 | 5 | 5.099019513592784 | 0.980486407216 E-3 | 4 E-3 |
| 26 | 1 | 5.1 | 4.898979485566356 | 01.02051433644 E-3 | 4 E-3 |
| 24 | -1 | 4.9 | 4.989990019950139 | 0.9980049861 E-4 | 4 E-4 |
| 25.1 | .1 | 5.01 | 4.9899909799498590 | 1.0020050141 E-4 | 4 E-4 |
| 24.9 | -.1 | 4.99 | 4.999999999999992 | 0.99998 E-6 | 4 E-6 |

I should mention another set of conventions that you will find in our book and often other places as well. To introduce it, we first write let \( h := x - a \). Then our expression becomes
\[
f(a + h) \sim f'(a)h + f(a),
\]
or better:
\[
f(a + h) - f(a) \sim f'(a)h
\]
This is useful because \( h \) is small, and the expression displays clearly how our approximation depends on this small number \( h \).

We could also write \( \Delta x \) in place of \( h \). Note that \( a \) could be anything, and could even be regarded as a “variable.” In fact to emphasize this, people tend
to write $x$ in place of $a$. Then our goal is to approximate $f(x + \Delta x) - f(x)$.

Here is the standard definition, using the language of “differentials,” in which we now write $dx$ in place of $\Delta x$.

**Definition:** Suppose $f$ is differentiable. For any $x$ in the domain of $f$ and any real number $dx$,

$$dy := f'(x)dx \quad \text{and} \quad \Delta y := f(x + dx) - f(x)$$

Our theorem then says that $dy$ is very near to $\Delta y$ if $dx$ is small. In fact, the difference between $dy$ and $\Delta y$ is small even compared to $dx$. 