STEEPEST DESCENT AND ASCENT

Math 225

The method of steepest descent is a numerical method for approximating local minima (and maxima) of differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$. The basic idea of the method is very simple: If the gradient is not zero where you are, then move in the direction opposite the gradient.

I find it easier to think in the opposite direction, namely, to seek out a local maximum by moving in the direction of the gradient. This could be called *steepest ascent.*

Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable. We want to approximate a point $\mathbf{a}$ where f takes a local maximum. Let $\mathbf{x}_0$ be an initial approximation (educated guess) of the maximum. The next approximation, $\mathbf{x}_1$, is obtained by adding a positive multiple of $\nabla f_{\mathbf{x}_0}$ to $\mathbf{x}_0$. This process is repeated, yielding a sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ in which

$$\mathbf{x}_{n+1} = \mathbf{x}_n + k\nabla f_{\mathbf{x}_n}.$$

If $\mathbf{x}_0$ was chosen close enough to $\mathbf{a}$, and if suitable conditions hold, then this sequence will converge to $\mathbf{a}$, that is, $\lim\limits_{n\to\infty} \mathbf{x}_n = \mathbf{a}$. (If you want to go towards a minimum, take k to be negative.)

A cleaner way to describe to describe this process is to define a mapping $T\colon \mathbb{R}^n \to \mathbb{R}^n$ by

$$(*) \qquad T(\mathbf{x}) = \mathbf{x} + k\nabla f_{\mathbf{x}}.$$

As a result we have that $\mathbf{x}_{n+1} = T(\mathbf{x}_n)$ for all $n = 0, 1, 2, \ldots,$ and so the sequence is obtained by repeated application of T to $\mathbf{x}_0$, *e.g.*, $\mathbf{x}_4 = T(T(T(T(\mathbf{x}_0))))$. The mapping T can clearly be used to generate a sequence starting with any point in $\mathbb{R}^n$—given a point in $\mathbb{R}^n$, the mapping T says where the next point is. One can think of T as describing the discrete motion of a particle as it jumps from one point of $\mathbb{R}^n$ to another. As such, it is an example of a discrete *dynamical system.*

Note that if $\mathbf{x} = \mathbf{a}$, the local maximum, then $\nabla f_{\mathbf{a}} = \mathbf{0}$ since $\mathbf{a}$ is a critical point. Thus

$$T(\mathbf{a}) = \mathbf{a} + k\nabla f_{\mathbf{a}} = \mathbf{a}.$$

Since $T(\mathbf{a}) = \mathbf{a}$, that is, T doesn't move $\mathbf{a}$, we call $\mathbf{a}$ a *fixed point* of T. Thus the problem of finding a minimum for f becomes one of finding a fixed point for T. Many important numerical methods are based on this idea of converting the problem at hand into a fixed point problem. Another method of this type that you may have seen is Newton's method.

The rate of convergence of $\mathbf{x}_n \to \mathbf{a}$ depends on the choice of k. If k is too small, the convergence is slow. On the other hand, if k is too large, the mapping can "jump" over the fixed point. Unfortunately, there isn't a single value of k that works best.

Whether or not a particular value of k works for a given function depends on the concavity of the function f—the greater the concavity the smaller k has to be. In fact, the value of k should really not be a constant at all! The value of k in $T(\mathbf{x}) = \mathbf{x} + k\nabla f_{\mathbf{x}}$ should depend on the concavity of f at $\mathbf{x}$, that is, the formula for T should really be

$$T(\mathbf{x}) = \mathbf{x} + \alpha(\mathbf{x})\nabla f_{\mathbf{x}},$$

where $\alpha\colon \mathbb{R}^n \to \mathbb{R}$ is some suitable positive function.

Here's the idea. Suppose we are at $\mathbf{x}$. We want to move in the direction of $\nabla f_{\mathbf{x}}$, but how far? Thinking of $\mathbf{x}$ as constant for the moment, consider the line through $\mathbf{x}$ in the direction of $\nabla f_{\mathbf{x}}$, namely $\gamma(t) = \mathbf{x} + t\nabla f_{\mathbf{x}}$. The Math 111 function obtained by restricting f to this line is

$$g(t) = f(\gamma(t)) = f(\mathbf{x} + t\nabla f_{\mathbf{x}}).$$

It seems that the appropriate place to move to in this direction is the point where g takes its maximum. It won't necessarily be the maximum for f, but it will be the best we can do in this direction.

For what value of t does g take its maximum? That's another approximation problem, but we are down to one variable, and so we can use Math 111/112 techniques. We will approximate g with its quadratic Taylor polynomial near $t = 0$:

$$g(t) \approx Q(t) = g(0) + g'(0)t + \frac{1}{2}g''(0)t^2.$$

The idea is that g will take its maximum at a value of t close to where Q takes its maximum. The graph of Q is a parabola, provided $g''(0) \neq 0$. For Q to have a maximum, its graph must be concave down. This is a reasonable expectation. If we are headed for a maximum of f, the graph of f is probably concave down, which would imply that the graphs of g and Q are concave down as well. Another thing to note is that $g'(0)$ is positive, since with increasing t we are moving in the direction of ∇f at $t = 0$.

You can easily verify that the value of t where Q takes its maximum under these assumptions is $t_* = -g'(0)/g''(0)$, and so we set

$$\alpha(\mathbf{x}) = t_* = -\frac{g'(0)}{g''(0)}.$$

Since $g'(0) > 0$ and $g''(0) < 0$, we have that $t_* > 0$, as expected.

We need formulas for $g'(0)$ and $g''(0)$ in terms of f. They are directional derivatives of f at $\mathbf{x}$ in the direction of $\nabla f_{\mathbf{x}}$. To simplify the notation, let $\mathbf{v} = \nabla f_{\mathbf{x}}$. Then

$$g'(0) = \frac{d}{dt}\bigg|_0 f(\mathbf{x} + t\mathbf{v}) = D_{\mathbf{v}}f(\mathbf{x}) = \nabla f_{\mathbf{x}} \cdot \mathbf{v} = \nabla f_{\mathbf{x}} \cdot \nabla f_{\mathbf{x}} = \|\nabla f_{\mathbf{x}}\|^2,$$

and

$$g''(0) = \frac{d^2}{dt^2}\bigg|_0 f(\mathbf{x} + t\mathbf{v}) = D^2_{\mathbf{v}} f(\mathbf{x}) = \mathbf{v} \cdot \big(Hf_{\mathbf{x}}\mathbf{v}\big) = \nabla f_{\mathbf{x}} \cdot \big(Hf_{\mathbf{x}}\nabla f_{\mathbf{x}}\big),$$

where $Hf_{\mathbf{x}}$ denotes the Hessian matrix of f at $\mathbf{x}$. Thus we have

$$\alpha(\mathbf{x}) = t_* = -\frac{\|\nabla f_{\mathbf{x}}\|^2}{\nabla f_{\mathbf{x}} \cdot \big(Hf_{\mathbf{x}}\nabla f_{\mathbf{x}}\big)}.$$

Plugging this expression into the formula for T we get

$$(**) \qquad T(\mathbf{x}) = \mathbf{x} - \frac{\|\nabla f_{\mathbf{x}}\|^2}{\nabla f_{\mathbf{x}} \cdot \big(Hf_{\mathbf{x}}\nabla f_{\mathbf{x}}\big)}\nabla f_{\mathbf{x}}.$$

Using $(**)$, as opposed to $(*)$, will result in much faster convergence to the maximum, but at a price. Like Newton's Method, this formula can be very sensitive to initial conditions. A poor choice $\mathbf{x}_0$ can result in movement away from the maximum. In fact, if you're near a minimum, you'll go there!

To see this, suppose $\mathbf{x}$ is in a region where f is concave up. Then the functions g and Q will also be concave up. The critical point of Q will be a minimum, and $\alpha(\mathbf{x}) = t_*$ will be negative instead of positive. Thus, in $(**)$ we will be going in the direction opposite the gradient, that is, towards a minimum.

Thus $(**)$ is both steepest ascent and descent depending on the initial approximation. In general, the map T can be chaotic. To use it to find a maximum, you should be sure you are in a region where f is concave down. Similarly, to find a minimum, you should be in a region where f is concave up. Otherwise, you should use the version with the constant coefficient. Note that with two or more variables, the graph of f can have *mixed* concavity! The conditions in the Second Derivative Test can be used to determine this.

Here are some functions to try ...

(1) $f(x, y) = -x^2 - 4y^2$
(2) $f(x, y) = -x^2 - y^4$
(3) $f(x, y) = x^2 - y^2$
(4) $f(x, y) = x^2 + xy + y^2$
(5) $f(x, y) = x^3 + 3xy - y^3$

Now for the challenge ... find the approximate values of the local minima of

$$f(x, y) = 1 + x - y - 2x^3 + x^6 + 6xy^2 + 3x^4y^2 + 3x^2y^4 + y^6$$

and their approximate locations. You should start by looking at the graph and level curves of f.

Robert Foote, Spring 1996, Revised Fall 2007