General discussion of 0th order methods.

We want to consider a class of zeroth order methods which <u>are</u> guaranteed to converge. First, we need to define convergence.

Definition. A numerical method M which produces a sequence $M(x) = x, x_1, \dots$ of points is said to be <u>globally convergent</u> for f if at least one of the limit points of $M(x)$ is a point where $\nabla f(x_*)$

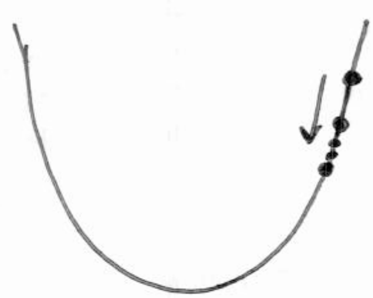
We are going to consider conditions under which a numerical method can fail to converge.

Note. $f(x_{k+1}) < f(x_k)$ is <u>not</u> enough to guarantee convergence.

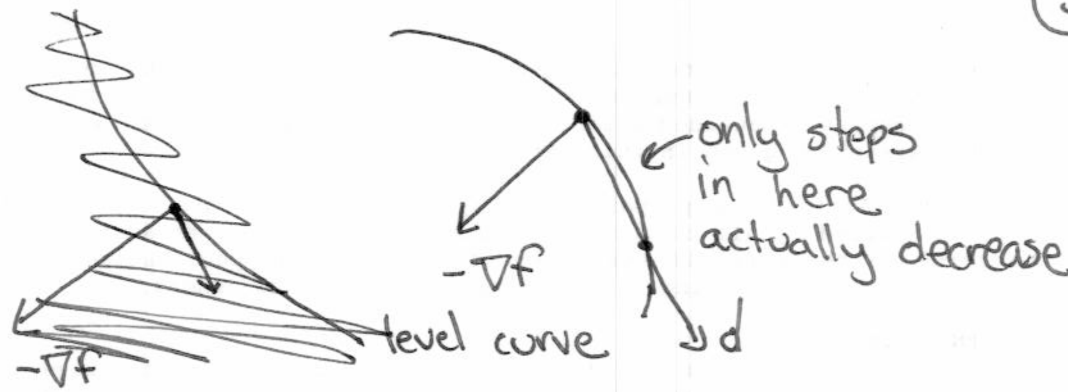
Examples.

a) $f(x) = x^2$, $x_k = (-1)^k (1 + 2^{-k})$

b) $f(x) = x^2$, $x_k = 1 + 2^{-k}$



c)



In each case, we have

Definition. d is a <u>descent direction</u> if

$$-\nabla f \cdot d > 0$$

and we have stepped in a descent direction. But

a) our steps were too long
b) our steps were too short
c) the descent directions → a non descent dir

We will show that if we arrange to avoid all these problems, our method must converge.

So consider Compass search.

Let Δ_k be our step size, Δ_0 an initial size, and Δ_{tol} a size beneath which we quit.

Step 1. Let $D_n = \{\pm e_1, \ldots, \pm e_n\}$ be the set of coordinate directions. Compute $f(x_k + \Delta_k d_k)$ for all $d_k \in D_n$.

Step 2. If $f(x_k + \Delta_k d_k) < f(x_k)$, for some d_k, ~~choose the smallest such~~ choose

$$x_{k+1} = x_k + \Delta_k d_k$$

$$\Delta_{k+1} = \Delta_k$$

Step 3. If not, the step fails

$$x_{k+1} = x_k$$

$$\Delta_{k+1} = \Delta_k / 2$$

If $\Delta_{k+1} < \Delta_{tol}$, terminate

The idea of proving that CS converges is pretty clever.

First, note that any vector in $\mathbb{R}^n$ makes an angle with $\cos\theta \geq 1/\sqrt{n}$ with some $d \in D_n$.

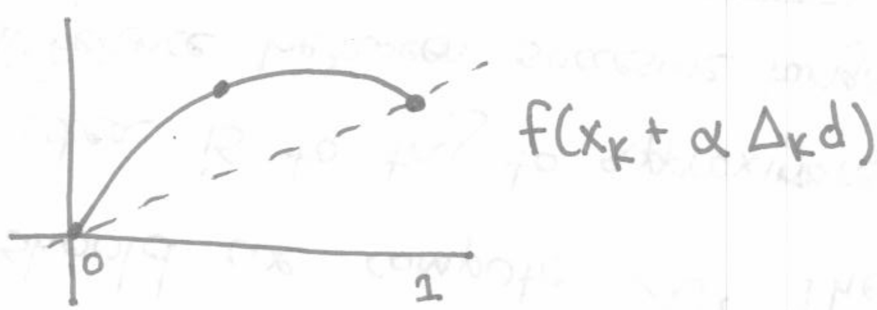
So

$$-\nabla f \cdot d = \|-\nabla f\| \, \|d\| \cos\theta$$
$$\geq \frac{1}{\sqrt{n}} \|\nabla f\| \, \|d\|.$$

Now if the step <u>failed</u>, we know

$$0 \leq f(x_k + \Delta_k d) - f(x_k)$$

Consider the function



By the mean value theorem, at some point $\alpha \in [0,1]$,

$$\frac{d}{d\alpha} f(x_k + \alpha \Delta_k d) = f(x_k + \Delta_k d) - f(x_k)$$

$$\parallel$$

$$\nabla f(x_k + \alpha \Delta_k d) \cdot \Delta_k d$$

So we have

$$\nabla f(x_k + \alpha \Delta_k d) \cdot \Delta_k d - \nabla f(x_k) \cdot \Delta_k d \geq -\nabla f(x_k) \cdot \Delta_k d$$

or

$$(\nabla f(x_k + \alpha \Delta_k d) - \nabla f(x_k)) \cdot \Delta_k d \geq -\nabla f(x_k) \cdot \Delta_k d$$

$$\geq \frac{1}{\sqrt{n}} \|\nabla f\| \, \|d\|$$

Now suppose ∇f is uniformly continuous. Then for some M

$$\| \nabla f(x_k + \alpha \Delta_k d) - \nabla f(x_k) \| \leq M \| \alpha \Delta_k d \| \leq$$

so

$$\leq M \| \Delta_k d \|$$

$$M \| \Delta_k d \| \geq \frac{1}{\sqrt{n}} \|\nabla f\| \, \|d\|$$

and

$$\|\nabla f\| \leq \sqrt{n}\, M \, \Delta_K$$

Now this is amazing! We have figured out a bound on $\|\nabla f\|$ at a failed step. Now if we could guarantee that

$\exists$ a sequence of failed steps of infinite length which converges

then we could conclude that $\nabla f = \vec{0}$ at the limit point because $\Delta_K \to 0$ as the # of failed steps goes up.