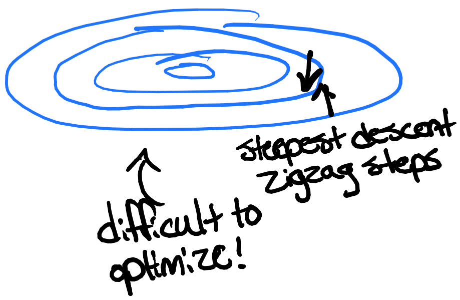
PHYS 177
winter 2019

lecture 17: Newton's method

learning goals
- apply Newton's method to optimization problems
- understand rates of convergence

last time, convergence was slow due to unbalanced curvature



we can overcome this problem by going to higher orders in perturbations

expand $f(\underline{x}+\underline{s})$ to second order

$$f(\underline{x}+\underline{s}) = f(\underline{x}) + \underline{s}^T \nabla f(\underline{x}) + \underline{s}^T \nabla^2 f(\underline{x}) \underline{s}$$

we can then find the direction $\underline{s}$ that minimizes $f(\underline{x}+\underline{s})$ by setting the derivative wrt. $\underline{s}$ to $\emptyset$!

$$\nabla f(\underline{x}) + \nabla^2 f(\underline{x}) \underline{s} = 0 \rightarrow \boxed{\underline{s} = -(\nabla^2 f(\underline{x}))^{-1} \nabla f(\underline{x})}$$

Newton step direction

rate of convergence

proofs are somewhat complicated, but for steepest descent we can show

$$f(\underline{x}_{k+1}) - f(\underline{x}^*) \leq \alpha [f(\underline{x}_k) - f(\underline{x}^*)]$$

linear approach to the minimum

note: for a quadratic function $f(\underline{x}) = \underline{x}^T \underline{\underline{a}}\, \underline{x} + \underline{b}^T \underline{x}$, $\alpha \propto \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2$, where $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ are eigenvalues of $\underline{\underline{a}}$

→ slow convergence when directions are unevenly scaled!

Newton's method instead converges quadratically,

$$|\nabla f(\underline{x}_{k+1})| \leq \alpha |\nabla f(\underline{x}_k)|^2 \quad , \text{ where } \alpha \propto [\nabla^2 f(\underline{x}^*)]^{-1} \text{ (convexity @ min)}$$

*notebook example w/ quadratic, Rosenbrock functions