**Jim Lambers**
**MAT 419/519**
**Summer Session 2011-12**
**Lecture 10 Notes**

These notes correspond to Section 3.2 in the text.

# The Method of Steepest Descent

When it is not possible to find the minimium of a function analytically, and therefore must use an iterative method for obtaining an approximate solution, Newton's Method can be an effective method, but it can also be unreliable. Therefore, we now consider another approach.

Given a function $f : \mathbb{R}^n \to \mathbb{R}$ that is differentiable at $\mathbf{x}_0$, the *direction of steepest descent* is the vector $-\nabla f(\mathbf{x}_0)$. To see this, consider the function

$$\varphi(t) = f(\mathbf{x}_0 + t\mathbf{u}),$$

where $\mathbf{u}$ is a *unit* vector; that is, $\|\mathbf{u}\| = 1$. Then, by the Chain Rule,

$$\begin{aligned} \varphi'(t) &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t} \\ &= \frac{\partial f}{\partial x_1}u_1 + \cdots + \frac{\partial f}{\partial x_n}u_n \\ &= \nabla f(\mathbf{x}_0 + t\mathbf{u}) \cdot \mathbf{u}, \end{aligned}$$

and therefore

$$\varphi'(0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{u} = \|\nabla f(\mathbf{x}_0)\| \cos\theta,$$

where $\theta$ is the angle between $\nabla f(\mathbf{x}_0)$ and $\mathbf{u}$. It follows that $\varphi'(0)$ is minimized when $\theta = \pi$, which yields

$$\mathbf{u} = -\frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|}, \quad \varphi'(0) = -\|\nabla f(\mathbf{x}_0)\|.$$

We can therefore reduce the problem of minimizing a function of several variables to a single-variable minimization problem, by finding the minimum of $\varphi(t)$ for this choice of $\mathbf{u}$. That is, we find the value of $t$, for $t > 0$, that minimizes

$$\varphi_0(t) = f(\mathbf{x}_0 - t\nabla f(\mathbf{x}_0)).$$

After finding the minimizer $t_0$, we can set

$$\mathbf{x}_1 = \mathbf{x}_0 - t_0\nabla f(\mathbf{x}_0)$$