Chapter 1
Several Gradients

These notes contain an introduction to the idea of Sobolev gradients and how they can be used in the study of differential equations. Numerical considerations are at once a motivation, an investigative tool and an application for this work.

First recall some facts about ordinary gradients. Suppose that for some positive integer \( n \), \( \phi \) is a real-valued \( C(1) \) function on \( \mathbb{R}^n \). It is customary to define the gradient \( \nabla \phi \) as the function on \( \mathbb{R}^n \) so that if \( x = (x_1, x_2, ..., x_n) \) is in \( \mathbb{R}^n \), then

\[
(\nabla \phi)(x) = \begin{pmatrix}
\phi_1(x_1, \cdots, x_n) \\
\vdots \\
\phi_n(x_1, \cdots, x_n)
\end{pmatrix}
\]  

(1.1)

where \( \phi_i(x_1, ..., x_n) \) is written in place of \( \partial \phi/\partial x_i, \ i = 1, 2, ..., n \).

The gradient \( \nabla \phi \) has the property that

\[
\lim_{t \to 0} \frac{1}{t} (\phi(x + th) - \phi(x)) = \phi'(x)h = \langle h, (\nabla \phi)(x) \rangle_{\mathbb{R}^n}, \ x, h \in \mathbb{R}^n,
\]  

(1.2)

and

\[
\| (\nabla \phi)(x) \|_{\mathbb{R}^n} = \sup_{h \in \mathbb{R}^n, \|h\|_{\mathbb{R}^n}=1} |\phi'(x)h|, \ x, h \in \mathbb{R}^n.
\]

Note that (1.2) can be taken as an equivalent definition of \( \nabla \phi \).

For \( \phi \) as above but with \( \langle \cdot, \cdot \rangle_S \) an inner product on \( \mathbb{R}^n \) different from the standard inner product \( \langle \cdot, \cdot \rangle_{\mathbb{R}^n} \), there is a function \( \nabla_S \phi : \mathbb{R}^n \to \mathbb{R}^n \) so that

\[
\phi'(x)h = \langle h, (\nabla_S \phi)(x) \rangle_S, \ x, h \in \mathbb{R}^n
\]

since the linear functional \( \phi'(x) \) can be represented using any inner product on \( \mathbb{R}^n \). Say that \( \nabla_S \phi \) is the gradient of \( \phi \) with respect to the inner product \( \langle \cdot, \cdot \rangle_S \) and note that the gradient \( \nabla_S \phi \) has properties similar to those of the ordinary gradient \( \nabla \phi \) above except for expression, (1.1).
From linear algebra, there is a linear transformation

\[ A : \mathbb{R}^n \to \mathbb{R}^n \]

which relates these two inner products in such a way that if \( x, y \in \mathbb{R}^n \), then

\[ (x, y)_S = (x, Ay)_{\mathbb{R}^n}. \]

Some reflection leads to

\[ (\nabla_S \phi)(x) = A^{-1}(\nabla \phi)(x), \quad x \in \mathbb{R}^n. \] (1.3)

Taking a cue from Riemannian geometry, one can have for each \( x \in \mathbb{R}^n \) an inner product

\[ \langle \cdot, \cdot \rangle_x \]

on \( \mathbb{R}^n \). That is, each point of \( \mathbb{R}^n \) can have its own inner product space. Consider such an assignment made together with a selection of a real-valued \( C^1 \) function \( \phi \) on \( \mathbb{R}^n \). Then for \( x \in \mathbb{R}^n \), define \( \nabla_x \phi : \mathbb{R}^n \to \mathbb{R}^n \) so that

\[ \phi'(x)h = \langle h, (\nabla_x \phi)(x) \rangle_x, \quad x, h \in \mathbb{R}^n. \]

For such a gradient system to be of much interest, the corresponding family of inner products, one inner product for each member of \( \mathbb{R}^n \), should be related to each other in an orderly way. This is similar to the case of Riemannian geometry in which it is required that inner products be assigned to tangent spaces in a differentiable fashion. In later chapters there are some natural assignments of inner product spaces, some related to Newton’s method, and some related to minimal surface problems.

Concrete aspects of the above discussion begin in the following chapter and continue throughout these notes. Most of these considerations apply to Hilbert spaces and, in a somewhat limited way, to more general spaces. Finite dimensional cases are for us synonymous with numerical considerations.

A central theme in these notes is that a given function \( \phi \) has a variety of gradients depending on choice of metric. More to the point, these various gradients have vastly different numerical and analytical properties even when arising from the same function. I first encountered the idea of variable metric in [174] where, in a descent process, different metrics are chosen as a process develops. Karmarkar [96] has used the idea with great success in a linear programming algorithm. In [104] and others, Karmarkar’s ideas are developed further. This writer has developed this idea (with differential equations in mind) in a series of papers starting in [145] (or maybe in [141]) and leading to [159, 161, 163]. Variable metrics are related to the conjugate gradient method [80]. Some other classical references to steepest descent are [38, 50, 208].
A ‘Sobolev gradient of $\phi$’ is a gradient of a $\phi$ when its domain is a finite or infinite dimensional Sobolev space.

There are two related versions of steepest descent. The earliest reference known to me for steepest descent is Cauchy [38]. The first version is discrete steepest descent, the second is continuous steepest descent.

Suppose one has an inner product $\langle \cdot, \cdot \rangle_S$ on a Hilbert space $H$, a real-valued $C^1$ function $\phi$ on $H$ and its gradient $\nabla_S \phi$. By ‘discrete steepest descent’ is meant an iterative process

$$x_n = x_{n-1} - \delta_{n-1}(\nabla_S \phi)(x_{n-1}), \ n = 1, 2, 3, ...,$$  \hspace{1cm} (1.4)

where $x_0$ is given and $\delta_{n-1}$ is chosen to be the number $\delta$ which minimizes, if possible,

$$\phi(x_{n-1} - \delta(\nabla_S \phi)(x_{n-1})), \ \delta \in \mathbb{R}.$$  

On the other hand, continuous steepest descent consists of finding a function $z : [0, \infty) \rightarrow H$ so that

$$z(0) = x \in H, \ z'(t) = -(\nabla_S \phi)(z(t)), \ t \geq 0.$$  \hspace{1cm} (1.5)

Continuous steepest descent may be interpreted as a limiting case of (1.4) in which, roughly speaking, various $\delta_n$ tend to zero (rather than being chosen optimally). Conversely, (1.4) might be considered (without the optimality condition on $\delta$) as a numerical method (Euler’s method) for approximating solutions to (1.5).

Using (1.4) one seeks $u = \lim_{n \rightarrow \infty} x_n$ so that

$$\phi(u) = 0$$  \hspace{1cm} (1.6)

or

$$\nabla_S \phi(u) = 0.$$  \hspace{1cm} (1.7)

Using (1.5) one seeks $u = \lim_{t \rightarrow \infty} z(t)$ so that (1.6) or (1.7) holds. Before more general forms of gradients are considered (for example where A in (1.3) is nonlinear), Chapter 2 gives an example intended to convince a reader that there are substantial issues concerning Sobolev gradients. It is hoped that Chapter 2 provides motivation for further reading even though later developments do not depend on proofs in Chapter 2. These arguments might be skipped in a first reading.

This introduction is closed with the indication of two applications of steepest descent:

(a) Many systems of differential equations have a variational principle, i.e. there is a function $\phi$ such that $u$ satisfies the system if and only if $u$ is a critical point of $\phi$. In such cases one tries to use steepest descent to find a zero of a gradient of $\phi$. 


(b) In other problems a system of nonlinear differential equations is written in the form
\[ F(x) = 0, \quad (1.8) \]
where \( F \) maps a Banach space \( H \) of functions into another such space \( K \). In some cases one might define for some \( p > 1 \), a function \( \phi : H \to \mathbb{R} \) by
\[ \phi(x) = \frac{1}{p} \|F(x)\|_H^p, \quad x \in H. \]
and then seek \( x \) satisfying (1.8) by means of steepest descent.

Problems of both kinds are considered. The following chapter contains an example of the second kind.
Sobolev Gradients and Differential Equations
Neuberger, J.
2010, XIII, 289 p., Softcover
ISBN: 978-3-642-04040-5