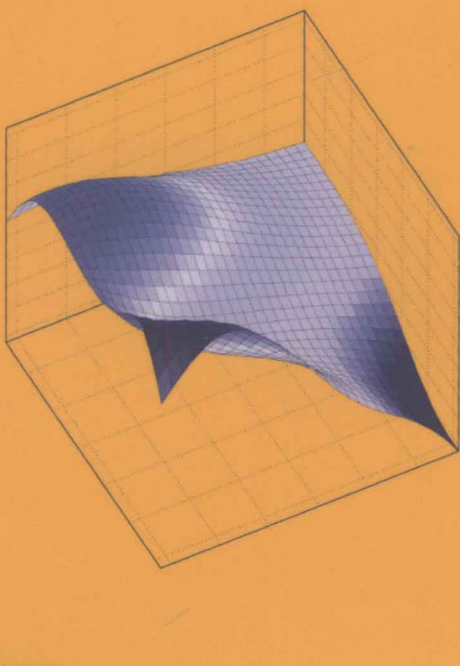
Lecture Notes in Mathematics

1670

J. W. Neuberger

Sobolev Gradients and Differential Equations

2nd Edition



Springer

J.W. Neuberger

Sobolev Gradients and Differential Equations

Second Edition

常州大学山书馆
藏书章

Springer

J.W. Neuberger
Department of Mathematics
University of North Texas
1155 Union Circle #311430
Denton, TX 76105-5017
USA
jwn@unt.edu

ISBN: 978-3-642-04040-5 e-ISBN: 978-3-642-04041-2
DOI: 10.1007/978-3-642-04041-2
Springer Heidelberg Dordrecht London New York

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2009938945

Mathematics Subject Classification (2000): 35A05, 35A15, 35A35, 65D10, 65D25, 46E35, 46T99, 34B15, 34B27, 31C25, 81V99

© Springer-Verlag Berlin Heidelberg 1997, 2010
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: SPi Publisher Services

Printed on acid-free paper

springer.com

Preface

What is expected from a theory of differential equations? Look first at the fundamental theorem for ordinary differential equations:

Theorem 0.1. *Suppose that* n *is a positive integer and* G *is an open subset of* $R \times R^n$ *which contains a point* (c, w). *Suppose also that* $f : G \to R^n$ *is a continuous function for which there is* $M > 0$ *such that*

$$\|f(t,x) - f(t,y)\| \leq M\|x - y\| \text{ for all } (t,x),(t,y) \in G. \tag{0.1}$$

Then there is an open interval (a, b) *containing* c *for which there is a unique function* u *on* (a, b) *so that*

$$u(c) = w, \ u'(t) = f(t, u(t)), \ t \in (a, b).$$

This result can be proved in several constructive ways which yield, along the way, error estimates giving a basis for numerical computation of solutions. Now this existence and uniqueness result certainly does not solve all problems in ordinary differential equations. For one thing, the result is only local. For just one other instance, it doesn't tell about two point boundary value problems, even though it has relevance there. Nevertheless, it provides a position of strength from which to study a wide variety of ordinary differential equations. The fact of existence and uniqueness of a solution gives us something to study in a qualitative, numerical or algebraic setting. The constructive nature of arguments for the above result gives one a good start toward discerning properties of solutions.

Many agree that it would be good to have a similar position of strength for partial differential equations but such does not now exist. It has been argued that there cannot be a central theory of partial differential equations since there is such a great variety of problems. To such an argument I reply that the same opinion about ordinary differential equations was probably held not so much more than a century ago.

These notes are devoted to a description of Sobolev gradients for a variety of problems in differential equations. Sobolev gradients are used in descent processes to find zeros or critical points of functions which in turn provide solutions to underlying differential equations. Our gradients are generally given constructively and do not require full boundary conditions (*i.e.*, conditions which are necessary and sufficient for existence and uniqueness) to be known beforehand. The processes tend to converge in some (non-Euclidean) sense to a nearest solution. The methods apply in cases which are mixed hyperbolic and elliptic — even cases in which regions of hyperbolicity and ellipticity are determined by nonlinearities. Applications to the problem of transonic flow will illustrate this. Numerics are a natural part of the development given here. In fact, numerics are in a sense ahead of theory, giving a spur to more inquiry.

So, do we arrive at a position of strength for fairly general partial differential equations? Here at least is a shadow of such a theory.

A key thing for a reader to keep in mind is that continuous steepest descent with Sobolev gradients is expressed as an ordinary differential equations in a function space whereas alternative descent methods are often partial differential equations themselves (for example, see Chapter 16 in the case of minimal surface problems).

Notes for Second Edition

The theory of Sobolev gradients has developed a great deal since the publication of the first edition of these notes. Many of these developments are reflected in this second edition, which is about twice the length of the first one.

- The use of Sobolev gradients to find critical points of the Ginzburg-Landau energy functional of superconductivity has greatly expanded. It is now near the design stage for superconducting devices. P. Kazemi's recent discoveries play a substantial role here.
- The treatment of Newton's method in the context of Sobolev gradients has been expanded to include a version of the Nash-Moser inverse function theorem. The problem of 'loss of derivatives' has been avoided entirely, a fact that leads to a relatively simple argument for such inverse function results when applied to differential equations. It was first pointed out by A. Castro that considerations for gradient inequalities have much in common with Moser's development of an inverse function theorem.
- The Tricomi equation, showing both elliptic and hyperbolic regions, has been treated using Sobolev gradients.
- A number of new convergence results for continuous steepest descent are included.
- Work on the hyperbolic Monge-Ampere equation, due to T. Howard, is described. This work opens up a new aspect of the study of such equations.
- Use of Sobolev gradients for nonlinear Schrödinger equations is noted.
- A greatly expanded list of properties of the imbedding operator which connects a Hilbert space with a dense linear subspace which is a Hilbert space in its own right. Much of this is due to P. Kazemi.

此为试读，需要完整PDF请访问：www.ertongbook.com

- After the first edition of this work was published, it was realized that this author's previous use of what is called 'gradient inequality' was preceded by Łojasiewicz inequalities in finite dimensions.
- There is reference to gradient inequality results work of S. Huang and of R. Chill.
- There is an account of Chan-Hilliard equations by S. Sial, T. Lookman, A. Saxena and the present writer.
- There are Sobolev gradient results for fractal regions.
- Some least squares results are given which have application to the problem of separating actual chaos from apparent chaos induced by discretization.
- A new result is given which relates nonlinear semigroup theory to the problem of boundary or supplementary conditions for partial differential equations.

In the first edition, several authors contributed sections on their work with Sobolev gradients. In the second edition, several have kindly agreed to write a chapter on their work. These include

- A development of numerical integration by means of Sobolev gradients, by Ian Knowles and Robert Wallace.
- A discussion of relationships between Sobolev gradients and preconditioning, by Janos Karatson.
- A presentation of curve fitting in the context of Sobolev gradients, by Robert Renka.
- Results on sign changing solutions and Morse index problems, by John M. Neuberger.
- Oil-water separation, elasticity and Model A problems, by Sultan Sial.

Robert Renka and I have had regular discussions about Sobolev gradients for more than two decades. Many others, particularly John M. Neuberger, have read portions of these notes and have contributed corrections and helpful suggestions. Any remaining errors and obscurities are mine. Many students, colleagues, collaborators and others have provided substantial insights. Any attempt at a list acknowledging this help would contain many names but would likely be inadequate. Hence I have decided to not try to make such a list.

I express profound gratitude to Springer for their help and extraordinary patience.

Contents

此为试读,需要完整PDF请访问: www.ertongbook.com

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, ϕ is a real-valued $C^{(1)}$ function on R^n. It is customary to define the gradient $\nabla\phi$ as the function on R^n so that if $x = (x_1, x_2, ..., x_n)$ is in R^n, then

$$(\nabla\phi)(x) = \begin{pmatrix} \phi_1(x_1, \cdots, x_n) \\ \vdots \\ \phi_n(x_1, \cdots, x_n) \end{pmatrix} \tag{1.1}$$

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

$$\lim_{t\to 0} \frac{1}{t}(\phi(x+th) - \phi(x)) = \phi'(x)h = \langle h, (\nabla\phi)(x)\rangle_{R^n}, \ x, h \in R^n, \tag{1.2}$$

and

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

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

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

$$\phi'(x)h = \langle h, (\nabla_S\phi)(x)\rangle_S, \ x, h \in R^n$$

since the linear functional $\phi'(x)$ can be represented using any inner product on R^n. Say that $\nabla_S\phi$ is the gradient of ϕ 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).

J.W. Neuberger, *Sobolev Gradients and Differential Equations*, Lecture Notes in Mathematics 1670, DOI 10.1007/978-3-642-04041-2_1,
© Springer-Verlag Berlin Heidelberg 2010

From linear algebra, there is a linear transformation

$$A : R^n \to R^n$$

which relates these two inner products in such a way that if $x, y \in R^n$, then

$$\langle x, y\rangle_S = \langle x, Ay\rangle_{R^n}.$$

Some reflection leads to

$$(\nabla_S \phi)(x) = A^{-1}(\nabla\phi)(x),\ x \in R^n. \tag{1.3}$$

Taking a cue from Riemannian geometry, one can have for each $x \in R^n$ an inner product

$$\langle\cdot,\cdot\rangle_x$$

on R^n. That is, each point of 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 ϕ on R^n. Then for $x \in R^n$, define $\nabla_x\phi : R^n \to R^n$ so that

$$\phi'(x)h = \langle h, (\nabla_x\phi)(x)\rangle_x,\ x, h \in R^n.$$

For such a gradient system to be of much interest, the corresponding family of inner products, one inner product for each member of 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 ϕ 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 ϕ' is a gradient of a ϕ 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 ϕ on H and its gradient $\nabla_S \phi$. By 'discrete steepest descent' is meant an iterative process

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

where x_0 is given and δ_{n-1} is chosen to be the number δ which minimizes, if possible,

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

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

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

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

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

$$\phi(u) = 0 \tag{1.6}$$

or

$$(\nabla_S \phi)(u) = 0. \tag{1.7}$$

Using (1.5) one seeks $u = \lim_{t \to \infty} z(t)$ 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 ϕ such that u satisfies the system if and only if u is a critical point of ϕ. In such cases one tries to use steepest descent to find a zero of a gradient of ϕ.

(b) In other problems a system of nonlinear differential equations is written in the form

$$F(x) = 0, \tag{1.8}$$

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 $\phi : H \to R$ by

$$\phi(x) = \frac{1}{p}\|F(x)\|_H^p, \ x \in H.$$

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.

Chapter 2
Comparison of Two Gradients

This chapter gives a comparison between conventional and Sobolev gradients for a finite dimensional problem associated with a simple differential equation. On first reading one might examine just enough to understand the statements of the two theorems. Nothing in the following chapters depends on the techniques of the proofs of these results. Although I expect similar theorems to exist for most systems of differential equations.

In this chapter, all norms and inner products which do not have a subscript are standard Euclidean.

Suppose that ϕ is a $C^{(2)}$ real-valued function on R^n and $\nabla_S \phi$ is the gradient associated with ϕ by means of the positive definite symmetric matrix A, as in the previous chapter. A measure of worth of $\nabla_S \phi$ in regard to a descent process is

$$\sup_{x \in R^n, \phi(x) \neq 0} \frac{\phi(x - \delta_x(\nabla_S \phi)(x))}{\phi(x)} \tag{2.1}$$

where, for each $x \in R^n$, $\delta_x \in R$ is chosen optimally, i.e. a number δ which minimizes

$$\phi(x - \delta(\nabla_S \phi)(x)), \ \delta > 0 \tag{2.2}$$

or, perhaps, is the least positive critical point of the above indicated function. Generally, the smaller the value in (2.1), the greater the worst case improvement in each discrete steepest descent step. It is remarked that $(\nabla_S \phi)(x)$ is a descent direction at x (unless $(\nabla_S \phi)(x) = 0$) since if

$$f(\delta) = \phi(x - \delta(\nabla_S \phi)(x)), \delta \geq 0,$$

then

$$f'(0) = -\|(\nabla_S \phi)(x)\|_S^2 < 0.$$

Equation (2.1) is used to compare performance of two gradients arising from the same function ϕ. For a simple example, choose ϕ so that

$$\phi(u) = u' - u \text{ on } [0,1], u \text{ absolutely continuous.} \tag{2.3}$$

J.W. Neuberger, *Sobolev Gradients and Differential Equations*, Lecture Notes in Mathematics 1670, DOI 10.1007/978-3-642-04041-2_2,
© Springer-Verlag Berlin Heidelberg 2010

For each positive integer n and with $\gamma_n = \frac{1}{n}$, define $\phi_n : R^{n+1} \to R$ so that if

$$x = (x_0, x_1, ..., x_n) \in R^{n+1},$$

then

$$\phi_n(x) = \frac{1}{2}\sum_{i=1}^{n}(\frac{x_i - x_{i-1}}{\gamma_n} - \frac{x_i + x_{i-1}}{2})^2. \tag{2.4}$$

Consider first the conventional gradient $\nabla\phi_n$ of ϕ_n. Pick $y \in C^{(3)}$ so that at least one of the following hold:

$$y'(0) - y(0) \neq 0,\ y'(1) - y(1) \neq 0. \tag{2.5}$$

Condition (2.5) amounts to the requirement that $y' - y$ **not** be in the domain of the adjoint of

$$L : Lz = z' - z,\ z \text{ absolutely continuous on } [0,1], \tag{2.6}$$

this adjoint being given by

$$L^t w = \{-(w' + w) : w \text{ absolutely continuous },w(0) = w(1) = 0\}, \tag{2.7}$$

(cf. [56]). Define a sequence of points $\{w^n\}_{n=1}^{\infty}$, $w^n \in R^{n+1}$, $n = 1, 2, ...,$ which are taken from y in the sense that for each positive integer n, w^n is the member of R^{n+1} so that

$$w_i^n = y(\frac{i}{n}),\ i = 0, 1, ..., n. \tag{2.8}$$

It will be shown that the measure of worth (2.1) deteriorates badly as the number of grid points approaches ∞. Specifically,

Theorem 2.1.

$$\lim_{n\to\infty} \phi_n(w^n - \delta_n \frac{(\nabla\phi_n)(w^n)}{\phi_n(w^n)}) = 1,$$

where for each positive integer n, δ_n is chosen optimally in the sense of (2.2).

This theorem expresses what many have seen in trying to use conventional steepest descent on differential equations. If one makes a definite choice for y with, say $y'(0) - y(0) \neq 0$, then one finds that the gradientsf $(\nabla\phi_n)(w^n)$, even for n quite small, have very large first component relative to all the others (except possibly the last one if $y'(1) - y(1) \neq 0$). This in itself renders $(\nabla\phi_n)(w^n)$ an unpromising object with which to perturb w^n in order that

$$w^{n+1} = w^n - \delta_n(\nabla\phi_n)(w^n)$$

此为试读,需要完整PDF请访问: www.ertongbook.com