

A CONVERGENCE RESULT FOR DISCREET STEEPEST DECENT IN WEIGHTED SOBOLEV SPACES

W. T. MAHAVIER

ABSTRACT. A convergence result is given for discrete descent based on Sobolev gradients arising from differential equations which may be expressed as quadratic forms. The argument is an extension of the result of David G. Luenberger on Euclidean descent and compliments the work of John W. Neuberger on Sobolev descent.

1. INTRODUCTION

Sobolev descent was introduced by J. W. Neuberger [7] and since that time, discrete Sobolev descent has been used with considerable success to solve specific problems [1], [8] and has been shown vastly superior to Euclidean descent in [5] and [6]. Neuberger proved convergence for general continuous settings and for a specific discrete problem in [6] and we prove a convergence result for the class of equations which may be expressed as a quadratic form which is symmetric and positive definite when restricted to the perturbation space arising from the boundary conditions of the differential equation under consideration. Note that this does not require that the gradient operator itself be symmetric. The proof generalizes the result of Luenberger [4] from discrete Euclidean descent to descent based on gradients arising from Sobolev inner products. Related methods such as generalized gradient methods and variable metric methods are treated in detail in [3]. An example is provided which indicates the improvements based on computing using a weighted Sobolev gradient versus Sobolev or Euclidean gradients on Legendre's equation.

1991 *Mathematics Subject Classification.* 65M12, 65N12, 65J05, 65K99.

Key words and phrases. Sobolev descent, convergence, generalized gradients, variable metrics.

Received: October 9, 1996.

©1996 Mancorp Publishing, Inc.

2. NOTATION AND STATEMENT OF THEOREM

Let $n \in \mathbf{N}$, $\langle \cdot, \cdot \rangle$ represent the Euclidean inner product, and E represent Euclidean $n+1$ -space, $(\mathfrak{R}^{n+1}, \langle \cdot, \cdot \rangle)$. Suppose $A \in L(E, E)$ is positive definite and symmetric. Define a (Sobolev) inner product by $\langle \cdot, \cdot \rangle_H = \langle A \cdot, \cdot \rangle$. Appropriate choices for A yield inner product spaces leading to vastly improved numerical results in terms of the number of iterations and time required to solve a given differential equation. See [5], [6], and the example in Section 4. Let H denote the space, $(\mathfrak{R}^{n+1}, \langle \cdot, \cdot \rangle_H)$ and H^0 the subspace of H which inherits the inner product and represents the perturbation space associated with the boundary conditions of the differential equation under consideration. If $B : H \rightarrow \mathfrak{R}$ is a linear operator on H representing the boundary conditions, let H^0 be the subspace of H defined by $H^0 = \{u \in H | Bu = 0\}$.

Suppose the functional representing the differential equation is given by

$$J(\cdot) = 1/2 \langle \cdot, Q \cdot \rangle$$

where $Q \in L(E, E)$ is symmetric and positive definite when restricted to H^0 .

Definition 1. *If S is any Hilbert space, $J : S \rightarrow \mathfrak{R}^+$ is a bounded linear functional, and for all $s \in S$ we define $\nabla_S J(s)$ to be the unique element in S such that $J'(s)(r) = \langle \nabla_S J(s), r \rangle_S$ for all $r \in S$.*

The Euclidean gradient, $(\nabla J)(u)$, the Sobolev gradient, $(\nabla_H J)(u)$, and the Sobolev gradient based on the perturbation space, $(\nabla_{H^0} J)(u)$, may now be defined based on this definition. If A was chosen based on the differential equation then $(\nabla_{H^0} J)$ now represents both the differential equation and the boundary conditions.

Theorem 1. *Suppose x^* is a minimizer of J and that x_0 and x^* satisfy both the boundary conditions and the condition, $\|x_0 - x^*\|_H < \|x_0 - y^*\|_H$ for any minimizer $y^* \neq x^*$, then $\lim_{k \rightarrow \infty} x_k = x^*$ where*

$$x_{k+1} = x_k - \delta_k (\nabla_{H^0} J)(x_k) \quad k = 0, 1, \dots$$

and δ_k minimizes $J(x_k - \delta_k (\nabla_{H^0} J)(x_k))$.

3. PROOF OF THEOREM

Lemma 1. *For $u \in E$, $\nabla J(u) = Qu$.*

Proof. The proof is by direct computation and is omitted. ■

Lemma 2. *(Neuberger) If P denotes the orthogonal projection of E onto H^0 under the Euclidean inner product, then for $u \in E$, $PA(\nabla_{H^0} J)(u) = P(\nabla J)(u)$.*

Proof. Given $u \in E$ we have for every $h \in H^0$,

$$\begin{aligned} \langle P(\nabla J)(u), h \rangle &= \langle (\nabla J)(u), h \rangle \\ &= J'(u)(h) \\ &= \langle (\nabla_{H^0} J)(u), h \rangle_H \\ &= \langle A(\nabla_{H^0} J)(u), h \rangle \\ &= \langle PA(\nabla_{H^0} J)(u), h \rangle. \blacksquare \end{aligned}$$

Lemma 3. *The gradient operator, $G = (\nabla_{H^0} J)$ is positive definite with respect to $\langle \cdot, \cdot \rangle_H$ when restricted to H^0 .*

Proof. Let $G = (\nabla_{H^0} J)$ and $x \in H^0$. Determining $y = Gx$ is equivalent to solving any of the following three systems over H^0 :

$$\begin{aligned} PAy &= PQx \\ (PA)|_{H^0}y &= (PQ)|_{H^0}x \\ A|_{H^0}y &= Q|_{H^0}x. \end{aligned}$$

Hence, $G = A|_{H^0}^{-1}Q|_{H^0}$ and

$$\langle x, A|_{H^0}^{-1}Q|_{H^0}x \rangle_H = \langle x, AA|_{H^0}^{-1}Q|_{H^0}x \rangle = \langle x, Q|_{H^0}x \rangle > 0. \blacksquare$$

The following lemma shows that a given function is bounded away from zero. In the case for the Euclidean inner product where the operator is symmetric, Kantorovich's inequality, [4], may be applied. However, bounds for the non-symmetric Sobolev case are not available.

Lemma 4. *If*

$$F(x) = \frac{\langle x, x \rangle_H \langle Gx, G^{-1}x \rangle_H}{\langle Gx, x \rangle_H \langle G^{-1}x, x \rangle_H}$$

then there exists a real number, c , such that for all $x \in H^0 - \{0\}$, $F(x) \geq c > 0$.

Proof. $\langle x, x \rangle_H$, $\langle Gx, x \rangle_H$, and $\langle G^{-1}x, x \rangle_H$ are clearly positive on $H^0 - \{0\}$. Since A is positive definite, we have

$$\begin{aligned} \langle Gx, G^{-1}x \rangle_H &= \langle A|_{H^0}^{-1}Q|_{H^0}x, Q|_{H^0}^{-1}A|_{H^0}x \rangle_H \\ &= \langle Q|_{H^0}x, Q|_{H^0}^{-1}A|_{H^0}x \rangle \\ &= \langle x, A|_{H^0}x \rangle > 0. \end{aligned}$$

If L parameterizes a line through the origin then direct computation shows that $F \circ L$ is constant on $\mathfrak{R}^{n+1} - \{0\}$. Suppose there exists a sequence p_k such that $F(p_k) \rightarrow 0$. After normalizing the sequence, there exists a convergent subsequence, q_k , and $F(q_k) \rightarrow 0$. Yet $q_k \rightarrow q$ implies $F(q_k) \rightarrow F(q)$, but $F(q) \neq 0$, a contradiction. We conclude, F is bounded away from zero on $H^0 - \{0\}$. \blacksquare

Lemma 5. *If $E(x) = \frac{1}{2} \langle x - x^*, G(x - x^*) \rangle_H$ and $g_k = Gx_k$ then*

$$E(x_{k+1}) = \left\{ 1 - \frac{\langle g_k, g_k \rangle_H \langle Gg_k, G^{-1}g_k \rangle_H}{\langle Gg_k, g_k \rangle_H \langle G^{-1}g_k, g_k \rangle_H} \right\} E(x_k)$$

Proof. The proof is a generalization of the argument in [4]. Putting $y_k = x_k - x^*$ and using the facts that $y_k = G^{-1}g_k$ and $Gx_k = Gy_k$ it may be shown that

$$\frac{E(x_k) - E(x_{k+1})}{E(x_k)} = \frac{\langle g_k, g_k \rangle_H \langle Gg_k, G^{-1}g_k \rangle_H}{\langle Gg_k, g_k \rangle_H \langle G^{-1}g_k, g_k \rangle_H}. \blacksquare$$

Hence, $E(x_k) \rightarrow 0$ and by the positive definite nature of $G|_{H^0}$ with respect to $\langle \cdot, \cdot \rangle_H$, $x_k \rightarrow x^*$ as desired.

4. EXAMPLE

As our example, we consider Legendre's equation. Suppose $Ku = 0$ where $Ku = ((1-t^2)u')' + 2u$ on $I = [0, 1]$ with $u(0) = 0$ (a forced initial condition), $u(1) = 1$, and $u \in C_I^2$. General solutions are $u(t) = c_1 t + \frac{c_2}{2} t \ln(\frac{1+t}{1-t})$ and only $u(t) = t$ satisfies the boundary conditions. The development of the continuous spaces, weighted derivatives, and generalized gradients on which the discrete theory is based provides motivation for the choices of the gradients and spaces considered here. A detailed treatment of the problem may be found in [5].

Denote the Euclidean norm by $\|\cdot\|$ and $x \in \mathfrak{R}^{n+1}$ by $x = (x_0, \dots, x_n)$. We assume no two consecutive components of x are zero. Suppose n is the number of divisions into which the interval $[0, 1]$ is partitioned and $\delta = 1/n$. Let $t_k = a + (k-1)\delta$ and $w_k = \sqrt{1-t_k^2}$ for all $k = 1, 2, \dots, n+1$.

Define discrete versions of the identity and weighted derivative operators, $D_0 : \mathfrak{R}^{n+1} \rightarrow \mathfrak{R}^n$, $D_1^w : \mathfrak{R}^{n+1} \rightarrow \mathfrak{R}^n$, and $D_w : \mathfrak{R}^{n+1} \rightarrow \mathfrak{R}^{2n}$ by

$$D_0(x) = \begin{pmatrix} \frac{x_1+x_2}{2} \\ \vdots \\ \frac{x_n+x_{n+1}}{2} \end{pmatrix}, \quad D_1^w(x) = \begin{pmatrix} \left(\frac{w_2+w_1}{2}\right) \left(\frac{x_2-x_1}{\delta}\right) \\ \vdots \\ \left(\frac{w_n+w_{n+1}}{2}\right) \left(\frac{x_{n+1}-x_n}{\delta}\right) \end{pmatrix}$$

and

$$D_w(x) = \begin{pmatrix} D_0(x) \\ D_1^w(x) \end{pmatrix}.$$

Define the unweighted derivative operator by $D_1 = D_1^w$, where w is the vector $(1, 1, \dots, 1)$.

We now consider descent based on the three spaces and refer to the three processes as Euclidean descent, Sobolev descent, and weighted Sobolev descent. The spaces are $(\mathfrak{R}^{n+1}, \langle \cdot, \cdot \rangle)$, $(\mathfrak{R}^{n+1}, \langle \cdot, \cdot \rangle_H)$, and $(\mathfrak{R}^{n+1}, \langle \cdot, \cdot \rangle_{H_w})$ where

$$\langle u, v \rangle_{H_w} = \langle D_0(u), D_0(v) \rangle + \langle D_1^w(u), D_1^w(v) \rangle$$

TABLE 1. Legendre's Equation

$(1 - t^2)y'' - 2ty' + 2y = 0 \quad y(0) = 0 \quad y(1) = 1 \quad N = 100$					
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	5948	24	10^{-6}	10^{-1}	6.6×10^{-1}
H	1998	7	10^{-6}	10^{-6}	3.7×10^{-5}
H_w	64	1	10^{-6}	10^{-7}	8.0×10^{-6}

$(1 - t^2)y'' - 2ty' + 2y = 0 \quad y(0) = 0 \quad y(1) = 1 \quad N = 10,000$					
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	2142	82	10^{-6}	10^{-6}	3.4×10^{-5}
H_w	85	3	10^{-6}	10^{-6}	1.2×10^{-5}

$(1 - t^2)y'' - 2ty' + 2y = 0 \quad y(0) = 0 \quad y(1) = 1 \quad N = 100,000$					
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	325	125	10^{-15}	10^{-14}	1.7×10^{-14}

for all $u, v \in \mathfrak{R}^{n+1}$. Notice that $\langle \cdot, \cdot \rangle_{H_w} = \langle \cdot, A \cdot \rangle$ where $A = D_0^t D_0 + (D_1^w)^t D_1^w$. To define the perturbation spaces associated with the equation, let $Bx = (x_0, x_n)$, $S = \{x \in \mathfrak{R}^{n+1} | Bx = 0\}$, $E^0 = E \cap S$, $H^0 = H \cap S$, and $H_w^0 = H_w \cap S$. Let $y \in \mathfrak{R}^{n+1}$, and define $J : (\mathfrak{R}^{n+1}, \|\cdot\|_{H_w}) \rightarrow \mathfrak{R}$ by

$$J(u) = \frac{1}{2} \sum_{k=1}^n \left(1 - \left(\frac{t_{k+1} + t_k}{2} \right)^2 \right) \left(\frac{u_{k+1} - u_k}{\delta} \right)^2 + \left(\frac{u_{k+1} + u_k}{2} \right)^2$$

Using Definition 1 to construct three gradients based on E^0 , H^0 , and H_w^0 the descent process is now,

$$x_{k+1} = x_k - \delta_k \nabla J(x_k) \quad k = 0, 1, \dots$$

where ∇ represents the Euclidean, Sobolev, or weighted Sobolev gradient and δ_k minimizes $J(x_k - \delta_k \nabla J(x_k))$. The algorithm is terminated when $\|y^{new} - y\| < \epsilon$ where y and y^{new} denote successive approximations to the solution. Table 1 represents the results based on these three descent processes.

5. CONCLUSIONS

Mathematicians and scientists have oft sought solutions to differential equations using descent based on the Euclidean gradient. The numerical work in this paper indicates that the choice of the underlying space and gradient are crucial for developing efficient numerical methods.

Boundary conditions are maintained at each step of the descent process guaranteeing exact boundary conditions for the solution and the method gives results on a small number of divisions which are representative of the results obtained on a large number of divisions making the method a candidate for multigrid techniques.

All work was performed on a NeXTstation 33 MHz 68040 Unix platform using the GNU C compiler. Codes for the problems and *Mathematica* codes for computing the necessary matrices are available from the author by e-mail at math-wtm@nich-nsunet.nich.edu.

REFERENCES

- [1] J. G. Dix and T. W. McCabe, *On finding equilibria for isotropic hyperelastic materials*, *Nonlinear Anal.* **15** (1990), 437–444.
- [2] J. Garza, *Using steepest descent to find energy-minimizing maps satisfying nonlinear constraints*, *Doct. Dissert.*, University of North Texas, 1994.
- [3] M. R. Hestenes, *Conjugate direction methods in optimization*, *Applications in Mathematics*, #12, Springer-Verlag, New York, 1980. 325pp.
- [4] D. G. Luenberger, *Linear and Nonlinear Programming* Addison-Wesley, Reading, 1989.
- [5] W. T. Mahavier, *A numerical method utilizing weighted Sobolev descent to solve singular differential equations*, *Nonl. World*, to appear.
- [6] J. W. Neuberger, *Steepest descent for general systems of linear differential equations in Hilbert space*, in “Ordinary Differential Equations and Operators”, *Lecture Notes in Math.*, Springer-Verlag, #1032, 1982, 390–406.
- [7] J. W. Neuberger, *Projection methods for linear and nonlinear systems of partial differential equations*, in “Ordinary Differential Equations and Operators”, *Lecture Notes in Math.*, Springer-Verlag, #546, 1976, 341–349.
- [8] R. J. Renka and J. W. Neuberger, *Minimal surfaces and Sobolev gradients*, *SIAM J. Sci. Comput.*, to appear.
- [9] G. Strang, *Introduction to applied mathematics*, Wellesley-Cambridge Press, Wellesley, 1986.

MATHEMATICS DEPARTMENT
NICHOLLS STATE UNIVERSITY
P.O. BOX 2026, THIBODAUX
LOUISIANA 70310, USA

E-mail address: math-wtm@nich-nsunet.nich.edu