

THE CONVEX ENVELOPE IS THE SOLUTION OF A NONLINEAR OBSTACLE PROBLEM

ADAM M. OBERMAN

(Communicated by Walter Craig)

ABSTRACT. We derive a nonlinear partial differential equation for the convex envelope of a given function. The solution is interpreted as the value function of an optimal stochastic control problem. The equation is solved numerically using a convergent finite difference scheme.

1. INTRODUCTION

We derive a partial differential equation in the form of a nonlinear obstacle problem for the convex envelope. While convex functions and the convex envelope have been the subject of study for many years, this equation requires the newer theory of viscosity solutions [5]. Interestingly, the result was anticipated about fifteen years ago by Griewank and Rabier [8], who observed that the convex envelope “appears [to be] a solution to a variant of the obstacle problem.”

The equation for the convex envelope, u , of the function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is

$$(Ob) \quad \max \{u(x) - g(x), -\lambda_1[u](x)\} = 0$$

(see Figure 1). Here $\lambda_1[u](x)$ is the smallest eigenvalue of the Hessian $D^2u(x)$.

Convexity is important in part because it links local and global properties of functions. For example, convexity is sufficient to ensure that a local minimum of a function is global. While the convex envelope is defined globally (see (CE) below), the partial differential equation (Ob) provides an entirely local characterization. Indeed, the proof requires Theorem 1, a local characterization of continuous convex functions.

Furthermore, the equation (Ob) allows the convex envelope to be interpreted as the value function of a stochastic control problem (see §3). This equation is solved numerically using a convergent finite difference scheme (see §4). Compared to existing numerical methods [3, 4, 10, 15], this scheme has the advantage that it is local and explicit. The local characterization is important for solving variational problems where convexity arises as a constraint (e.g., [14]), one of the motivations for this work.

In dimensions larger than one, (Ob) is an obstacle problem involving a fully nonlinear partial differential equation. This equation is degenerate elliptic in the

Received by the editors November 29, 2005.

2000 *Mathematics Subject Classification*. Primary 35J70, 52A41; Secondary 93E20, 65N06.

Key words and phrases. Convex envelope, obstacle problem, partial differential equation.

It is a pleasure to acknowledge Luis Silvestre for valuable discussions.

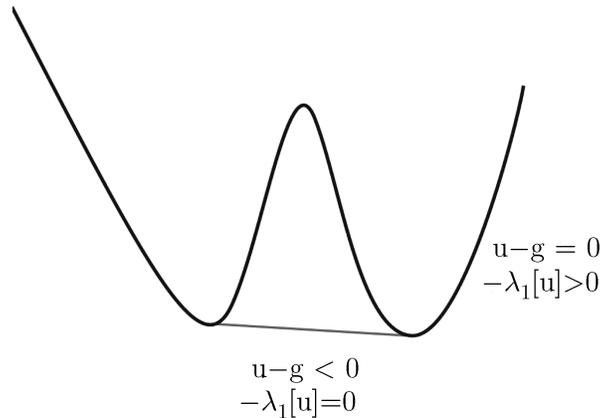


FIGURE 1. Illustration of the equation

sense of Crandall-Lions [5]: it is increasing in u , and decreasing in the Hessian of u . In fact, the equation (Ob) is a combination of examples from the introduction of the User's Guide: obstacle problems [5, ex. 1.7], and functions of the eigenvalues [5, ex. 1.8].

Known results regarding the regularity of the convex envelope apply to solutions of (Ob). While certainly convex functions need not be differentiable in general, whenever the function g is continuously differentiable, so too is the convex envelope [9].

2. DERIVATION OF THE EQUATION

Begin by recalling the characterization of the convex envelope of the function $g(x)$ as the supremum of all the convex functions which are majorized by g ,

$$(CE) \quad u(x) = \sup\{v(x) \mid v \text{ convex, } v(y) \leq g(y) \text{ for all } y \in \mathbb{R}^n\}.$$

Next recall the construction of viscosity solutions by Perron's method,

$$(Pe) \quad u(x) = \sup\{v(x) \mid v(x) \text{ is a subsolution of (Ob)}\}.$$

It is well known that twice differentiable convex functions are characterized by the property that the Hessian is everywhere positive semidefinite [2, Prop 1.2.6]. This result is generalized to continuous functions by the following theorem, whose proof is postponed to the end of the section.

Theorem 1. *The continuous function $u : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if it is a viscosity solution of $-\lambda_1[u] \leq 0$.*

Definition 1. The upper semicontinuous function u is a viscosity solution of $-\lambda_1[u] \leq 0$ if for every twice-differentiable function $\phi(x)$,

$$(1) \quad -\lambda_1[\phi](x) \leq 0, \quad \text{whenever } x \text{ is a local maximum of } u - \phi.$$

Next we define viscosity solutions of (Ob) in terms of sub- and supersolutions: a function is a viscosity solution of (Ob) if it is both a subsolution and a supersolution.

Definition 2. The upper semicontinuous function u is a viscosity subsolution of (Ob) if for every twice-differentiable function $\phi(x)$,

$$\phi(x) - g(x) \leq 0 \quad \text{and} \quad -\lambda_1[\phi(x)] \leq 0, \quad \text{whenever } x \text{ is a local maximum of } u - \phi.$$

The lower semicontinuous function u is a viscosity supersolution of (Ob) if for every twice-differentiable function $\phi(x)$,

$$\phi(x) - g(x) \geq 0 \quad \text{or} \quad -\lambda_1[\phi(x)] \geq 0, \quad \text{whenever } x \text{ is a local minimum of } u - \phi.$$

Theorem 2. *The convex envelope of the function g is a viscosity solution of (Ob).*

Proof. By Theorem 1 and Definition 2, subsolutions of (Ob) consist precisely of those convex functions majorized by g . So the supremums in the definition of the convex envelope (CE) and the Perron formula (Pe) for the solution of (Ob) are over the same sets, and therefore the functions are equal. \square

Write $x = (x^1, x')$ for $x = (x^1, x^2, \dots, x^n)$ in \mathbb{R}^n . For a given set B in \mathbb{R}^n , define the cylindrical neighborhood, $C_\epsilon(B)$, by

$$C_\epsilon(B) \equiv \{y \in \mathbb{R}^n \mid |y^1 - x^1| \leq \epsilon \text{ and } \|y' - x'\| \leq \epsilon \text{ for some } x = (x^1, x') \in B\}.$$

Proof of Theorem 1. 1. Suppose u is convex. Let x_0 be a local maximum of $u - \phi$ for some twice-differentiable function ϕ . We can assume $u(x_0) = \phi(x_0)$. Let $p(x)$ be a supporting hyperplane for u at x_0 . Then x_0 is a local maximum of $p - \phi$, so $D^2\phi(x_0) \geq D^2p(x_0) = 0$, which gives (1).

2. Suppose $u(x)$ is not convex. Then there are points $w, y, z \in \mathbb{R}^n$ with $w = \theta y + (1 - \theta)z$, for $0 < \theta < 1$, such that

$$(2) \quad u(w) > \theta u(y) + (1 - \theta)u(z).$$

Choose coordinates so that w, y, z lie on the x^1 axis, and $y^1 < z^1$. Let $q(x) = a(x^1 - w^1)^2/2 + b(x^1 - w^1) + u(w)$ be the parabola which interpolates u at the points w, y, z . By assumption (2), $a < 0$.

(i) We begin with the case $n = 1$. If w is a local maximum of $u - q$, then we are done. If not, let

$$x_* \in \operatorname{argmax}\{u(x) - q(x) \mid x \in [y, z]\}.$$

Then $x_* \in (y, z)$; therefore, x_* is a local maximum of $u - q$, and we have obtained a contradiction to (1) since $\lambda_1[q] = a < 0$.

(ii) Now suppose $n > 1$. Redefine $q(x) = (a + \delta)(x^1 - w^1)^2/2 + b(x^1 - w^1) + u(w)$, for $0 < \delta < -a$. Then $q(w) = u(w)$, but $q > u$ at y, z . Thus $q > u$ in a neighborhood of each of the points y, z . Choose $\epsilon > 0$ so that

$$q(x) > u(x), \quad \text{for } x \in C_\epsilon(y) \cup C_\epsilon(z).$$

Define $\phi(x) = q(x) + M\|x'\|^2/\epsilon^2$ where $M = \max\{u(x) \mid x \in C_\epsilon(I)\}$ and I is the line segment $[y^1 - \epsilon, z^1 + \epsilon]$. Then $\lambda_1[\phi] = a - \delta < 0$, $\phi(w) = u(w)$. Finally, $\phi \geq u$ on $\partial C_\epsilon(I)$ since $q \geq u$ when $\|x'\| < \epsilon$ and $M \geq u$ when $\|x'\| = \epsilon$. If w is a local maximum of $u - \phi$, then we are done. If not, choose

$$x_* \in \operatorname{argmax}\{u(x) - \phi(x) \mid x \in C_\epsilon(I)\}.$$

Then x_* is in the interior of $C_\epsilon(I)$, so x_* is a local maximum of $u - \phi$ and we have obtained a contradiction to (1). \square

3. VARIATIONAL AND STOCHASTIC CONTROL INTERPRETATIONS

The classical obstacle problem is a model for the equilibrium position of a membrane that is restricted to remain above (or below, in our case) the obstacle whose height is represented by the function $g(x)$. The solution, u , is the minimizer of the Dirichlet integral, $\int |Du|^2$, subject to the pointwise constraints $u \leq g$. It satisfies the following variational inequality (see [6, pp. 467–9]),

$$(3) \quad \max\{u(x) - g(x), -\Delta u(x)\} = 0.$$

Notice that in one dimension, (3) and (Ob) coincide.

The variational inequality (3) can be recast as the Hamilton-Jacobi-Bellman (HJB) equation for a stochastic optimal stopping problem. While a natural variational interpretation of (Ob) seems unlikely, it also has an interpretation as the value function of a stochastic control problem.

Next we will apply the Dynamic Programming Principle (DPP) to derive the HJB equations (3), (Ob). For clarity, our derivation is formal; we refer to [7] for a rigorous derivation of related equations.

3.1. Optimal stopping interpretation of the obstacle problem. Consider the diffusion corresponding to a particle undergoing Brownian motion in \mathbb{R}^n ,

$$(4) \quad \begin{cases} d\mathbf{x}(t) = \sqrt{2} d\mathbf{w}(t), \\ \mathbf{x}(0) = x_0, \end{cases}$$

(we use boldface to indicate a diffusion process in \mathbb{R}^n). Suppose that at a time t we can stop the process and incur a cost of $g(\mathbf{x}(t))$. The objective is to choose a stopping time, $\tau(\cdot)$, to minimize the expected cost

$$J(x_0, \tau) \equiv E[g(\mathbf{x}(\tau)) \mid \mathbf{x}(0) = x_0] \equiv E^{x_0}[g(\mathbf{x}(\tau))].$$

The DPP begins by generalizing the problem to consider different starting points. It then relates the optimal cost at these different points. So define the *value function*,

$$u(x) \equiv \min_{\tau} J(x, \tau),$$

where the minimization is over all stopping times.

Next, assume the value function is known for points near x_0 and consider the following near-optimal strategy. One option is to stop immediately, which costs $g(x_0)$. An alternative is to let the diffusion proceed for time t and thereafter follow the optimal strategy, which costs $E^{x_0}[u(\mathbf{x}(s))]$. Choosing the least of these two alternatives gives

$$u(x_0) = \min \{g(x_0), E^{x_0}[u(\mathbf{x}(t))]\} + o(t).$$

This part of the argument is formal: in fact we have only obtained an inequality. See [7] for the additional argument required to obtain the reverse inequality. Subtract to obtain

$$\max \{u(x_0) - g(x_0), u(x_0) - E^{x_0}[u(\mathbf{x}(t))]\} = o(t).$$

The infinitesimal generator corresponding to the diffusion (4) is simply the Laplacian (see, e.g., [13]), so we have

$$(5) \quad \lim_{t \rightarrow 0} \frac{E^x[u(\mathbf{x}(t))] - u(x)}{t} = \Delta u(x).$$

Using (5) and taking the limit $t \rightarrow 0$ gives (3), since x_0 was arbitrary.

3.2. **Stochastic control interpretation for the convex envelope.** Consider the diffusion

$$(6) \quad \begin{cases} d\mathbf{x}(t) = \sqrt{2}\theta(t)dw(t), \\ \mathbf{x}(0) = x_0, \end{cases}$$

where now w is a one-dimensional Brownian motion, and the control, $\theta(\cdot)$, is a mapping into unit vectors in \mathbb{R}^n . The objective is to minimize the expected cost

$$J(x, \theta(\cdot), \tau) \equiv E^x[g(x_\tau)]$$

over all stopping times, $\tau(\cdot)$, and direction paths $\theta(\cdot)$. The value function is

$$u(x) = \min_{\theta(\cdot), \tau} J(x, \theta(\cdot), \tau).$$

As before, we formally apply the DPP. One strategy is to stop the process; this costs $g(x_0)$. Another strategy is to fix $\theta(\cdot) = \theta$ to be constant, and let the diffusion proceed for time t , thereafter following the optimal strategy. This strategy costs $E^{x_0, \theta}[u(\mathbf{x}(s))] \equiv E^{x_0}[u(\mathbf{x}(s)) \mid \theta(\cdot) = \theta]$. Minimizing over θ , and over the two strategies gives

$$u(x_0) = \min \left\{ g(x_0), \min_{\theta} E^{x_0, \theta}[u(\mathbf{x}(t))] \right\} + o(t).$$

Subtract to get

$$\max \left\{ u(x_0) - g(x_0), u(x_0) - \min_{\theta} E^{x_0, \theta}[u(\mathbf{x}(t))] \right\} = o(t).$$

Using the definition of infinitesimal generator corresponding to the diffusion (4) with θ fixed (see, e.g., [13]) gives

$$(7) \quad \lim_{t \rightarrow 0} \frac{E^{x, \theta}[u(\mathbf{x}(t))] - u(x)}{t} = \theta^T D^2 u(x) \theta.$$

Using (7) in the preceding equation and taking the limit $t \rightarrow 0$ yields

$$\max \left\{ u(x) - g(x), - \inf_{|\theta|=1} \theta^T D^2 u(x) \theta \right\} = 0.$$

Finally, using the Courant-Hilbert characterization of the eigenvalues recovers (Ob).

4. NUMERICAL SOLUTION OF THE EQUATION

The following is a semi-discrete, monotone approximation of $\lambda_1[u](x)$:

$$\lambda_1^\epsilon[u](x) \equiv \min_{|v|=1} \frac{u(x + \epsilon v) - 2u(x) + u(x - \epsilon v)}{\epsilon^2} = \lambda_1[u](x) + O(\epsilon^2).$$

This yields the natural approximation to (Ob), $\max \{u(x) - g(x), -\lambda_1^\epsilon[u](x)\} = 0$, which is monotone and consistent [12], and therefore convergent [1]. The resulting nonlinear equation can be solved iteratively [12].

To fully discretize $\lambda_1[u]$, replace the minimum over $|v| = 1$ by a minimum over a finite number of grid directions, using a wide stencil scheme [11]. A solution is displayed in Figure 2.

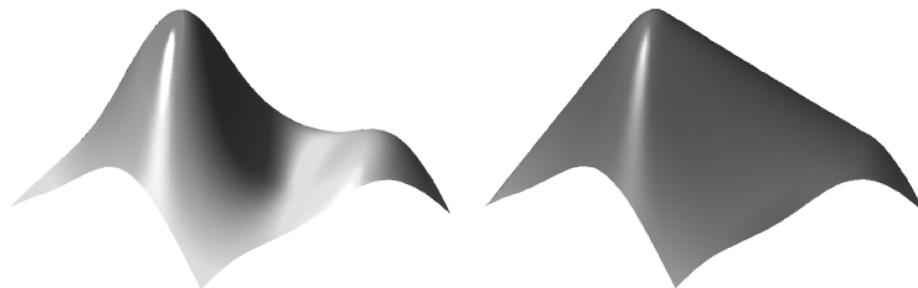


FIGURE 2. The function g and its convex envelope (inverted).

REFERENCES

1. Guy Barles and Panagiotis E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, *Asymptotic Anal.* **4** (1991), no. 3, 271–283. MR1115933 (92d:35137)
2. Dimitri P. Bertsekas, *Convex analysis and optimization*, Athena Scientific, Belmont, MA, 2003, With Angelia Nedić and Asuman E. Ozdaglar. MR2184037 (2006j:90001)
3. Yann Brenier, *Un algorithme rapide pour le calcul de transformées de Legendre-Fenchel discrètes*, *C. R. Acad. Sci. Paris Sér. I Math.* **308** (1989), no. 20, 587–589. MR1001813 (90f:65242)
4. Bernard Brighi and Michel Chipot, *Approximated convex envelope of a function*, *SIAM J. Numer. Anal.* **31** (1994), no. 1, 128–148. MR1259969 (94m:49049)
5. Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User's guide to viscosity solutions of second order partial differential equations*, *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), no. 1, 1–67. MR1118699 (92j:35050)
6. Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845 (99e:35001)
7. Wendell H. Fleming and H. Mete Soner, *Controlled Markov processes and viscosity solutions*, Applications of Mathematics (New York), vol. 25, Springer-Verlag, New York, 1993. MR1199811 (94e:93004)
8. A. Griewank and P. J. Rabier, *On the smoothness of convex envelopes*, *Trans. Amer. Math. Soc.* **322** (1990), no. 2, 691–709. MR986024 (91k:49021)
9. Bernd Kirchheim and Jan Kristensen, *Differentiability of convex envelopes*, *C. R. Acad. Sci. Paris Sér. I Math.* **333** (2001), no. 8, 725–728. MR1868942 (2002g:49024)
10. Y. Lucet, *A fast computational algorithm for the Legendre-Fenchel transform*, *Comput. Optim. Appl.* **6** (1996), no. 1, 27–57. MR1394296 (98i:90066)
11. Adam M. Oberman, *Convergent difference schemes for functions of the eigenvalues*, submitted.
12. ———, *Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems*, *SIAM J. Numer. Anal.* **44** (2006), no. 2, 879–895 (electronic). MR2218974 (2007a:65173)
13. Bernt Øksendal, *Stochastic differential equations*, sixth ed., Universitext, Springer-Verlag, Berlin, 2003, An introduction with applications. MR2001996 (2004e:60102)
14. J.C. Rochet and P. Choné, *Ironing, sweeping and multidimensional screening*, *Econometrica* **66** (1998), 783–826.
15. Luminata Vese, *A method to convexify functions via curve evolution*, *Comm. Partial Differential Equations* **24** (1999), no. 9-10, 1573–1591. MR1708102 (2000f:35066)

DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, BURNABY, BRITISH COLUMBIA, CANADA V5A 1S6

E-mail address: aoberman@sfu.ca