The Calculus of Variations

April 23, 2007

The lectures focused on the Calculus of Variations. As part of optimization theory, the Calculus of Variations originated in 1696 when Johann Bernoulli posed the brachistochrone problem. This problem related to the curve between two points along which a ball would require minimal time of travel to reach the bottom. This problem was solved by many different mathematicians of the time, some of which included Newton, l'Hôpital, and Bernoulli.

On February 9, 16, and 23, 2007, Professor Andrej Cherkaev, graduate student Russ Richins, and Professor David Dobson presented the basic ideas and some practical applications of the Calculus of Variations. Transcribing these lectures were John Tate, Jianyu Wang and Yimin Chen.

Professor Andrej Cherkaev's Lecture:

Introduction:

In its early days, the calculus of variations had a strong geometric flavor. Later, new problems arose during the 19th century, and the calculus of variations became more prevalent and was used to solve problems such as finding minimal surfaces.

The name "calculus of variations" emerged in the 19th century. Originally it came from representing a perturbed curve using a Taylor polynomial plus some other term, and this additional term was called the variation. Thus, the variation is analogous to a differential, in which the rate of change of a function at a point can be used to estimate the function value at a nearby point. Now mathematicians more commonly use minimizing sequences instead of variations, but the name variation has remained.

Mathematical Techniques:

A simple problem to illustrate the spirit of the calculus of variations is as follows: Given a function f, for $x \in \mathbb{R}$, minimize $f(x)$ over x with the following conditions:

$$\frac{\delta f}{\delta x} = 0,$$

$$\frac{\delta^2 f}{\delta x^2} \geq 0.$$

Thus we use the Second Derivative Test to find the relative minima of $f(x)$, and then we order the minima to find the absolute minimum of the function.

However, certain problems arise. For example, if $f(x)$ is not differentiable, then $f'(x)$ would be be discontinuous. In calculus of variations, the function is usually assumed to be differentiable. Nonetheless, sometimes the derivative might not exist or might go to infinity. Other complications might include the nonexistence of a solution or a function with multiple minima of equal value, such as $f(x) = (x^2 - 1)^2$. In this case, because the value of the function at $x = 1$ and $x = -1$ is the same, the function has no unique absolute minimum.

Two ways to approach these difficulties using the calculus of variations are called regularization and relaxation.

Regularization:

Regularization at the most general level consists of adding something to function to be minimized to make it more regular. For example, consider the aforementioned function $f(x) = |x|$. An example of regularization would be to set this equal to

$$\lim_{n\to\infty} f_n(x),$$

where

$$f_n(x) = \sqrt{x^2 + \frac{1}{n^2}}.$$

These functions have unique minima which converge to the minimum of f.

The regularization thus eliminated the problem of a discontinuous first derivative at $x = 0$.

One common type of problems in which regularization is used are viscosity problems in which sometimes the derivative is not continuous, but the calculus of variations used with regularization allows a solution to be found which is almost optimal.

In summary, regularization corrects the weird behavior of a function, allowing one to find a solution.

Relaxation:

Relaxation, at the most general level, includes adding extra conditions to a problem to relax the constraints. For example, analytically an open set of differentiable functions might be given, but we would like to use an optimal closed set or perhaps even the closure of some open set. Another example of relaxation might include functions for which boundary conditions include inequalities, and relaxation would include modifying some of those inequalities.

Let us illustrate the use of regularization and relaxation by some example problems.

A Problem Using Regularization: Minimize the integral $F = \int_{-1}^{1} x^2 u'^2 dx$ over the set of functions u satisfying the boundary conditions $u(-1) = -1$ and $u(1) = 1$.

Here we will apply one of the basic tools of the Calculus of Variations, the Euler-Lagrange Equation. This equation is as follows: Given that there exists a twice-differentiable function $u = u(x)$ satisfying the conditions $u(x_1) = u_1$ and $u(x_2) = u_2$ which minimizes an integral of the form

$$I = \int_{x_1}^{x_2} f(x, u, u')dx,$$

what is the differential equation for $u(x)$?

Stated in terms of our problem, $x_1 = -1$ and $x_2 = 1$, and $f(x, u, u') = x^2 u'^2$. We will now use the Euler-Lagrange differential equation, which is

$$\frac{\delta f}{\delta u} - \frac{d}{dx}\left(\frac{\delta f}{\delta u'}\right) = 0.$$

After computing the derivatives, we find that

$$\frac{d}{dx}(2x^2 u') = 0,$$

which means $2x^2 u' = c$, some constant. Thus $u' = \frac{c}{2x^2}$, and $u = \frac{-C}{2x} + D$ for constants C and D. But then u is discontinuous for $x = 0$, and so substituting $x^2 + \epsilon^2$ for x^2 we consider the regularized problem

$$I_2 = \min_u \int (x^2 + \epsilon^2) u'^2 dx.$$

Then

$$u' = \frac{c}{2(x^2 + \epsilon^2)},$$

and $u' = \frac{c}{x^2+\epsilon^2}$. Then we find that

$$u = c \times \arctan(\frac{x}{\epsilon}).$$

Addressing the boundary conditions that $u(-1) = -1$ and $u(1) = 1$, as ϵ tends to zero, we find that the constant is $\frac{1}{\arctan(\frac{x}{\epsilon})}$. For $x > 0$,

$$\lim_{\epsilon \to 0} \frac{1}{\arctan(\frac{x}{\epsilon})} = \frac{\pi}{2},$$

and for $x < 0$,

$$\lim_{\epsilon \to 0} \frac{1}{\arctan(\frac{x}{\epsilon})} = \frac{-\pi}{2}.$$

Thus we have determined that our function u is $\frac{\pi}{2}\arctan\left(\frac{x}{\epsilon}\right)$ for $x > 0$ and $-\frac{\pi}{2}\arctan\left(\frac{x}{\epsilon}\right)$ for $x < 0$.

A Problem Using Relaxation: Minimize over u $\int_0^1((u'^2-1)^2+u^2)dx$ subject to the boundary condition $u(0) = u(1) = 1$.

The minimum corresponds to $u = 0$ and $u = \pm 1$. We can relax this problem by considering $\hat{u} : |\hat{u}'| = 1$. The solution thus could be a function that oscillates infinitely often, with the derivative's graph like a square wave with amplitude 1.

Applications of Calculus of Variations:

Since its inception, the calculus of variations has been applied to a variety of problems. In engineering, it has been used for optimal design. Previously, much engineering was done through trial and error. For example, many cathedrals toppled during their construction, and parts of them had to be repeatedly rebuilt before stability was achieved. In fact, until the late 17th century, much of engineering design was a matter of building structures, modifying them in the case of failure rather than knowing beforehand whether or not they were adequately strong. In mechanics, the calculus of variations allows one to find the stability of a system mathematically without actually building it. One such modern use is a number of Java-based bridge modeling applications which are available online. In physics and materials science, the calculus of variations has been applied to the study of composite materials. Finally, a new use of this subject has been in evolutionary biology, in which the calculus of variations is used to explore optimization in nature.

In all of these fields, the application of the calculus of variations is often design when the goal is not known, such as the design of a bridge when the exact loads and locations of those loads at any given time are not known. Mathematically, the calculus of variations relates to bifurcation problems and problems such as the maximization of a minimal eigenvalue.

Graduate Study in the Calculus of Variations at Utah:

In the Spring 2007 semester, a 5000-level class on the Calculus of Variations is being taught. Beginning in the Fall 2007 semester, a 6000-level class on the topic will become regular. Nonetheless, other classes cover material concerned with the calculus of variations. Some of those classes include a topics class on Optimization Methods, Partial Differential Equations, and Applied Mathematics (depending on the instructor).

Concluding Quote

"The calculus of variations is so much in the heart of differential equations and analysis that you cannot get away from it."

Graduate Student Russ Richins' lecture:

Given a domain Ω in $\mathbb{R}^2$, and $u : \Omega \longrightarrow \mathbb{R}$, a continuous function, then the surface area of the graph of u is given by

$$\int_\Omega \sqrt{1+|\nabla u|^2}dx.$$

One of the basic problems in the Calculus of Variations is the so-called minimizing surface area problem, that is to solve the minimization problem

$$\inf_{u\in\mathcal{A}} \int_\Omega \sqrt{1+|\nabla u|^2}dx, \quad \mathcal{A} = \{u \in C^1(\Omega) : u|_{\partial\Omega} = g\},$$

where g is a function defined on the boundary of Ω.

A more general problem is to find

$$\inf_{u\in\mathcal{A}} \int_\Omega f(x,u,\nabla u)dx,$$

where f is a continuous function, and $\mathcal{A}$ is a given set of functions.

Now look at variational problems in a more abstract setting. Given an admissible set of functions $\mathcal{A}$, find u that solves the problem

$$I(u) = \inf_{v\in\mathcal{A}} I(v),$$

where $I : \mathcal{A} \longrightarrow \overline{\mathbb{R}}$ is a given continuous functional.

So we are looking for

$$\inf_{v\in\mathcal{A}} I(v).$$

Since we are seeking an infimum, it must be that there exists a sequence $\{v_k\} \subset \mathcal{A}$ such that

$$\lim_{k\to\infty} I(v_k) = \inf_{v\in\mathcal{A}} I(v).$$

Such a sequence $\{v_k\}$ is called a minimizing sequence.

In Calculus classes, we learn to solve similar optimization problems, where $\mathcal{A}$ is a closed interval and I is a continuous function. The requirement of compactness and continuity are sufficient to guarantee that the solution exists in this case. In the general framework, compactness and continuity will again be essential. The following theorem states conditions for the existence of solutions to the main problem in the Calculus of Variations.

Theorem 1 *Let $\Omega \subset \mathbb{R}^n$ be a bounded, open set with Lipschitz boundary. Let $f \in C(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n)$ satisfy:*
(a) The map $\xi \mapsto f(x, u, \xi)$ is convex for every $(x, u) \in \overline{\Omega} \times \mathbb{R}$.
(b) There exists $p > q \geq 1$ and $\alpha_1 > 0$, α_2, $\alpha_3 \in \mathbb{R}$, such that

$$f(x, u, \xi) \geq \alpha_1 |\xi|^p + \alpha_2 |u|^q + \alpha_3,$$

$\forall (x, u, \xi) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n$.

Then if $u_0 \in W^{1,p}(\Omega)$ with $I(u_0) < \infty$, there exists $u \in u_0 + W_0^{1,p}(\Omega) =: \mathcal{A}$, such that

$$I(u) = \inf_{v \in \mathcal{A}} I(v) = \inf_{v \in \mathcal{A}} \int_\Omega f(x, v(x), \nabla v(x)) dx.$$

Furthermore, if $(u, \xi) \mapsto f(x, u, \xi)$ is strictly convex for every $x \in \overline{\Omega}$, then the minimizer is unique.

In order to be clear, the space $W^{1,p}(\Omega)$ is a subset of $L^p(\Omega)$ whose elements are functions which are differentiable in a certain sense. We say that $v \in L^p(\Omega)$ has a weak first partial derivative with respect to x_i, if there exists a function $g \in L^p(\Omega)$ such that

$$\int_\Omega v \frac{\partial \varphi}{\partial x_i} dx = - \int_\Omega g \varphi dx, \qquad \forall \varphi \in C_0^\infty(\Omega).$$

The space $W^{1,p}(\Omega)$ consists of all functions in $L^p(\Omega)$ with weak first partial derivatives in all directions and is a Banach Space, when endowed with the norm

$$\| v \|_{W^{1,p}(\Omega)} = \{ \| v \|^p_{L^p(\Omega)} + \| \nabla v \|^p_{L^p(\Omega)} \}^{1/p}.$$

The following two examples illustrate why the conditions in the above theorem are optimal.

***Example* 1 :** Consider the function $f(u, \xi) = \sqrt{u^2 + \xi^2}$, find the minimizer, if it exists.

Solution: In this case, condition (b) in the theorem holds only with $p = 1$.We define

$$\mathcal{A} = \{ v \in W^{1,1}(0,1) : v(0) = 0, \quad v(1) = 1 \}, \quad m = \inf_{v \in \mathcal{A}} I(v),$$

then

$$I(u) \geq \int_0^1 \mid u'(x) \mid dx \geq \int_0^1 u'(x) dx = u(1) - u(0) = 1, \quad \forall u \in \mathcal{A}.$$

Therefore $m \geq 1$.

We now construct a minimizing sequence v_k as follows: For $k = 1, 2, 3, \cdots$, v_k is given by

$$v_k(x) = \begin{cases} 0 & if \quad x \in [0, 1 - 1/k] \\ 1 + k(x - 1) & if \quad x \in (1 - 1/k, 1] \end{cases},$$

then $I(v_k) = \int_{1-1/k}^{1} \sqrt{(1+k(x-1))^2 + k^2}dx \le \frac{1}{k}\sqrt{1+k^2} \to 1$, as $k \to \infty$.

Therefore $m = 1$.

Assume that there exists a minimizer $u \in \mathcal{A}$, then

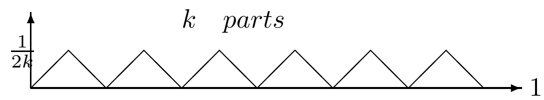
$$1 = I(u) = \int_0^1 \sqrt{u^2 + u'^2}dx \ge \int_0^1 \mid u' \mid dx \ge \int_0^1 u'dx = u(1) - u(0) = 1.$$

This implies that all equalities hold in the line above, and therefore $u = 0$ almost everywhere in $(0,1)$, which in turn implies that $u' = 0$ almost everywhere in $(0,1)$. The minimizer could only be the zero function, but it does not satisfy the boundary conditions. Therefore, there can be no minimizer for this problem.

***Example* 2 :** Let $f(u,\xi) = (\xi^2 - 1)^2 + u^4$, $\mathcal{A} = \{v \in W_0^{1,4}(0,1)\}$, $m = \inf_{v\in\mathcal{A}} I(v)$. Show $m = 0$.

Solution: Again, we construct a minimizing sequence v_k. For each integer ν with $0 \le \nu \le k-1$, where $k \ge 2$, let $\{v_k\}$ be defined on $[\frac{\nu}{k}, \frac{\nu+1}{k}]$ as follows:

$$v_k(x) = \begin{cases} x - \frac{\nu}{k} & if \quad x \in [\frac{\nu}{k}, \frac{2\nu+1}{2k}] \\ -x + \frac{\nu+1}{k} & if \quad x \in (\frac{2\nu+1}{2k}, \frac{\nu+1}{k}] \end{cases} .$$



Then $\mid v_k' \mid = 1$ almost everywhere, and $\mid v_k \mid \le \frac{1}{2k}$, so $0 \le I(v_k) \le \frac{1}{(2k)^4} \to 0$, as $k \to \infty$. Therefore $m = 0$. But if $u \in W_0^{1,4}(0,1)$ is a minimizer, then $I(u) = 0$, and $u = 0$, $\mid u' \mid = 1$ almost everywhere in $(0,1)$. It ends up that the elements of $W_0^{1,4}(0,1)$ are continuous, so this is impossible.

This example explains again that the minimizer does not always exist. The reason for this is that f is not convex. In the first example, when $p = 1$, $W^{1,p}(\Omega)$ is not reflexive.

One of the very important properties needed to hold for functionals to have a minimizer is "Weak Lower Semi Continuity", which is defined by: If $\{V_k\}$ converges weakly to v, then

$$I(v) \le \liminf_{k\to\infty} I(V_k).$$

Professor David Dobson's lecture:

Introduction:
Given a symmetric, positive definite matrix $A \in \mathbb{R}^{n\times n}$,

$$A(x) = (a_{ij}(x)), x \in \mathbb{R},\ a_{ij}(x) \in C^{\infty},$$

assume $A(x)$ has eigenvalues $0 \le \lambda_1(x) \le \lambda_2(x) \le \cdots \le \lambda_n(x)$. We want to find $Sup_{x\in[0,1]}\lambda_1(x)$.

A possible way to do this is using the freshman approach: compute $\lambda_1'(x)$ explicitly, and find x_0 which satisfies $\lambda_1'(x_0) = 0$. Therefore the solution will be at $x = 0, x = x_0 \in (0, 1)$ or $x = 1$. But the problem is that sometimes $\lambda_1'(x)$ does not exist.

For example, if

$$A(x) = \begin{pmatrix} 1+x^2 & 0 \\ 0 & 1+(x-1)^2 \end{pmatrix},$$

then λ_1 is $1 + x^2$ when $x \in [0, \frac{1}{2}]$ and is $1 + (x - 1)^2$ when $x \in [\frac{1}{2}, 1]$. It fails to be differentiable at $x = \frac{1}{2}$. So sometimes we cannot ignore the situations when the eigenvalue functions are not differentiable.

Buckling Load problem:
Another interesting example is to find the shape of the strongest column. Lagrange posed the problem: which curve determines the column of unit length and unit volume with maximum resistance to buckling under exact compression.

Keller's solution:
In 1960, Keller used the Euler buckling load

$$\lambda_1 = \inf_{u\in V} \frac{\int_0^1 EI|u|''^2 dx}{\int_0^1 |u|''^2 dx},$$

here, $u(x)$ is the displacement of the column, E is Young's modulus, $I(x)$ is given by $CA^2(x)$, and $A(x)$ is the cross sectional area and $V \subset \mathbb{H}^2(0, 1)$, which incorporates the boundary conditions, $u(0) = u'(0) = u(1) = u(1) = 0$.

If λ is an eigenvalue of the matrix A with corresponding eigenvector u. Then $\langle u, Au\rangle = \lambda\langle u, u\rangle$. So $\lambda = \frac{\langle u,Au\rangle}{\langle u,u\rangle}$. This is exactly the Raleigh Quotient. Then the smallest eigenvalue λ_1 is the minimizer of the Raleigh Quotient. And u is the solution of $-(EIu'')'' = \lambda_1 u$, subject to the boundary conditions $u(0) = u'(0) = u(1) = u'(1) = 0$.

Keller's solution gives cross sectional area $A(x) = 0$ at $x = \frac{1}{4}, x = \frac{3}{4}$.

Cox's solution:
In 1992, Cox Overton showed: (a) The above solution is not correct. (b) The buckling load calculated by Keller is 16 times greater than the actual buckling load.

So what went wrong? The answer is that Keller assumed λ_1 is a simple eigenvalue, while it is not. We can find the correct solution by numerical methods using techniques of nonsmooth analysis by Clarke.

Conclusion:
For the eigenvalues' optimization problems, we may not assume eigenvalues are simple. There are some cases where we need to find other possible methods to solve those problems.