Calculus of Variations, Lecture 8, 3/26/2013

1st half: optimization with a pde constraint (the "adjoint eqn" method) then another look at the Pontryagin Max Principle

2nd half: start multidimensional calc of vars, by discussing nonlinear elasticity

Consider the following question: for $a \in \mathbb{R}$, let $u = u_a$ solve

$$-\Delta u + a u = 1 \quad \text{in } D$$
$$u = 0 \quad \text{at } \partial D$$

where D is a fixed domain. Problem: find

$$\min_{a \in \mathbb{R}} \int_D |u_a - g|^2,$$

or more generally evaluate $f'(a)$ where

$$F(a) = \int_D |u_a - g|^2.$$

Why are we asking this?

1) Optimal control is pretty similar, except constraint is an ODE instead of a pde + the control is a function of s rather than constant

2) There are lots of situations where we're interested in pde's with parameters, + we may want to optimize some functional of the solution, or to know how it depends on the parameters.

Answer to the problem is this: given any $a \in \mathbb{R}$, we have

$$F'(a) = -2\int_D u_a p_a$$

where $p = p_a$ solves $\quad -\Delta p + ap = u_a - g$

Explanation: since $F(a) = \int_D |u_a - g|^2$ we have

$$F'(a) = \int_D 2(u_a - g)\,\dot{u}_a$$

where $\dot{u}_a = \frac{d}{da} u_a$. But differentiating the eqn. gives

$$-\Delta \dot{u}_a + u_a + a\dot{u}_a = 0 \quad \text{in } D.$$
$$\dot{u}_a = 0 \quad \text{at } \partial D$$

Now, multiply eqn for $p = p_a$ by $\dot{u}_a$ to get

$$\int_D (u_a - g)\dot{u}_a = \int_D (-\Delta p_a + a\, p_a)\, \dot{u}_a$$

$$= \int_D (-\Delta \dot{u}_a + a \dot{u}_a)\, p_a$$

$$= -\int_D u_a\, p_a$$

We can derive the Pontr Max Prin by a very similar argument. The main differences, as noted earlier, are that the control is a fn and the eqn is an ode.

Consider, as in lecture 7, $\max_{\alpha(s)} \int_t^T h(y(s),\alpha(s))\,ds + g(y(T))$

with state eqn $y_s = f(y,\alpha)$. We want to take 1st varn of the objective, so for any $\alpha(s)$ and $\dot{\alpha}(s)$ let

$$\alpha(s,\varepsilon) = \alpha(s) + \varepsilon\, \dot{\alpha}(s)$$

$$\dot{y}(s) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} y(s,\varepsilon)\quad, \text{etc.}$$

The optimality condition is of course that the

1st variation of the objective vanishes for all choices of $\dot{\alpha}(s)$ (I assume there are no ptwise restrictions on $\alpha(s)$), ie

$$\int_t^T h_y \dot{y} + h_\alpha \dot{\alpha} \, ds + g_y(y(T))\, \dot{y}(T) = 0$$

for all $\dot{\alpha}(s)$. Differentiating the state eqn gives

$$\dot{y}_s = f_y \dot{y} + f_\alpha \dot{\alpha} \quad , \quad \dot{y}(t) = 0 .$$

Making an inspired choice, we let $p(s)$ solve

$$p_s = -f_y(y(s),\alpha(s))\, p(s) - h_y(y(s),\alpha(s)) \quad t<s<T$$

$$p(T) = g_y(y(T))$$

(secretly: $p(s)$ will be the $\pi(s)$ of the PMP, and 1st eqn says $p_s = -\nabla_y H$)

Then

$$\int_t^T p_s \dot{y} = \int_t^T -f_y \dot{y} p - h_y \dot{y}$$

$$LHS = g_y(y(T))\,\dot{y}(T) - \cancelto{0}{p(t)\dot{y}(t)} - \int_t^T p \dot{y}_s$$

$$RHS = \int_t^T -p\dot{y}_s + f_\alpha p \dot{\alpha} - h_y \dot{y}$$

So $$g_y(y(T))\,\dot{y}(T) = \int_t^T f_\alpha p \dot{\alpha} - h_y \dot{y}$$

Thus: 1st varn of objective is

$$\int_t^T h_y \dot{y} + h_\alpha \dot{\alpha}\, ds + g_y(y(T))\,\dot{y}(T)$$

$$= \int_t^T (h_\alpha + f_\alpha p)\,\dot{\alpha}(s)\, ds.$$

1st varn $= 0$ for all $\dot{\alpha}(s)$ $\iff$ $h_\alpha + f_\alpha p = 0$

$\iff$ $\alpha(s)$ achieves the optimum at $y(s), p(s)$ of $H(p,x) = \max_\alpha \{ f(x,\alpha)\cdot p + h(x,\alpha) \}$

So: from this perspective $p(s)$ (formerly $\pi(s)$) solves the "adjoint eqn", and the cond that α be optimal for defn of H is the optimality condition.

Advantage of this viewpt: if you have guessed $\alpha(s)$ and it isn't optimal, we can get a better choice by "steepest descent," simply taking $\dot{\alpha}(s) = -(h_\alpha + f_\alpha p)$ along the path assoc to $\alpha(s)$.

Exercise related to the above:

(1) Returning to the calculation at the beginning of the section, suppose the pde is

$$-\Delta u + a x u_x + b y u_y = 1 \quad \text{in } D$$
$$u = 0 \quad \text{at } \partial D$$

where $D \subset \mathbb{R}^2$ and $a, b \in \mathbb{R}$. Consider once again

$$F(a,b) = \int_D |u - g|^2 .$$

Use the "adjoint eqn" method to obtain expressions for $\partial F/\partial a$ and $\partial F/\partial b$.

(2) Now let's let the unknown coefficient be a function of x: given $a(x)$, defined on D, let $u = u_a$ solve

$$-\Delta u + a(x)\, u = 1 \quad \text{in } D$$
$$u = 0 \quad \text{at } \partial D$$

and consider

$$F[a] = \int_D |u_a - g|^2$$

where $g(x)$ is given. If $a(x,\varepsilon) = a(x) + \varepsilon \dot{a}(x)$, find an expression for $\dot{F} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} F[a(x,\varepsilon)]$ that's of the form

$$\dot{F} = \int_D \varphi(x)\, \dot{a}(x)\, dx$$

where φ depends on $a(x)$ but not on $\dot{a}(x)$. (φ is, essentially, the "gradient" of F with respect to $a(x)$.)

We now shift focus toward <u>multidimensional problems</u> in the calculus of variations.

Recall our discussion (early in the semester) of existence: we showed that it's suff't (for pbm of form $\min \int W(Du)$ to have a soln) that W be convex, with L^p growth at ∞.

In the scalar setting ($u: \mathbb{R}^n \to \mathbb{R}$) convexity is also a necessary condition for a robust existence theory (we'll discuss this soon). But for $u: \mathbb{R}^n \to \mathbb{R}^m$ ($n>1$, $m>1$) convexity is not really the right condn (it is suff't but not necessary).

Moreover this <u>matters</u>, eg for elasticity. (Good sources on this: Dacorogna's book, on reserve; or Ball's 1977 article in Arch Rational Mech + Anal.)

The var'l pbm assoc elasticity is:

$$\min_{u|_{\partial\Omega}=u_0} \int_\Omega W(F) \, dx$$

where Ω = reference (undeformed body) $\subset \mathbb{R}^3$
(assumed stress-free)

$x: \Omega \to \mathbb{R}^3$ deformation, takes $X \in \Omega$ to
$x(X)$ = deformed position

$F = \frac{\partial x}{\partial X}$ — $n \times n$ matrix ("deformation gradient")

$W(RF) = W(F)$ for all $R \in SO(3)$.
("frame indifference").

Last condn $\Leftrightarrow$ $W(F)$ depends only on $(F^TF)^{1/2}$, since by polar decompn

$$F = \underset{\text{rotn}}{R} \; \underset{\text{pos def symm}}{(F^TF)^{1/2}}$$

[eigs of $(F^TF)^{1/2}$ are the "principal stretches"]

whenever F is a 3x3 orientation-preserving lin map. (There's also an obvious analogue of all this for "2D elasticity".)

A bit more abt W:

a) $F = I$ is a local min (since undeformed state is stress-free)

b) $W \to \infty$ as $\det F \to 0$ and as $|F| \to \infty$ (large compression or extension should be difficult)

c) if material is isotropic then $W(F) = W(FR)$ for all rtns R; corresp to $W(F)$ being a symmetric fn of the eigenvalues of $(F^T F)^{1/2}$ ("principal stretches")

Key pt for us: W cannot be a convex fn of F, since $SO(3)$ is not convex.

In fact, we expect W to be minimized at both

$$I \text{ and } R_0 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

but $W = \infty$ at $\frac{1}{2}I + \frac{1}{2}R_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$!

So: for plms of elasticity we need a class of nonconvex var'l plms where

(1) minimizer is achieved

(2) preceding properties of W are possible

For simplicity let's focus on analogous issue in 2D (so $x(X) : \mathbb{R}^2 \to \mathbb{R}^2$ + F is 2×2; everything discussed above still applies).

Ball showed that the class of "polyconvex fns" has the desired properties (1) + (2)

In 2D, $W(F)$ is polyconvex (by defn) if it has the form

$$W(F) = \varphi(F, \det F)$$ with φ a convex fn of its 5 arguments.

<u>Key questions</u>:

(i) is this class large enough to include reasonable elastic energies?

(ii) why is $\int W(F)$ lower semicontinuous when $W = \varphi(F, \det F)$ with φ convex?

Answer to (i): yes! The choice

$$W(F) = A(v_1^\alpha + v_2^\alpha - c) + g(\det F)$$

works, where $v_1 + v_2$ are eigs of $(F^TF)^{1/2}$ (so $\det F = v_1 v_2$) and g is convex; here $A + c$ are constants, g must be convex, + the restrn $W = \min$ at $F = I$ places some restriction on them.

Polyconvexity of such W is due to

<u>Lemma</u>: a symmetric fn of v_1, v_2 determines a <u>convex</u> fn of F if it is convex + increasing in each v_i

(Pf of Lemma is not trivial; see eg Ball's 1977 paper.)

Answer to (ii): key pt is that <u>det Dx is wkly cont's</u> under wk convergence in $W^{1,p}$ for $p > 2$.

More careful stmt: in this setting of maps $\mathbb{R}^2 \to \mathbb{R}^2$, $\det(\partial x/\partial X)$ is quadratic in $\partial x_i/\partial X_\alpha$. So if a sequence $x^\nu(X)$ is bdd in $W^{1,p}$, $\det(\partial x^\nu/\partial X)$ is bdd in $L^{p/2}$. If $p > 2$, $p/2 > 1$, so this seq is precompact in <u>wk topology</u> on $L^{p/2}$. So (for a subsequence) $\det(\partial x^\nu/\partial X)$ converges wkly to <u>some</u> $L^{p/2}$ fn. Our assertion is that the wk limit is $\det(\partial x^\infty/\partial X)$ (if $x^\nu \to x^\infty$ wkly in $W^{1,p}$).

If we grant the assertion above then existence via direct method is easy. Need only prove lower semicontinuity.

$\underline{1^{st}\text{ pf}}$ use Fenchel transform of φ, where $W = \varphi(F, \det F)$.

Evidently $\varphi(F, \det F) = \sup_{G, t} \; G\cdot F + t\cdot \det F - \varphi^*(G,t)$

So

$$\int \varphi(Dx, \det Dx) = \sup_{G(x), t(x)} \int G\cdot F(x) + t\cdot \det F - \varphi^*(G,t).$$

= sup of wkly conts fns

$\underline{2^{nd}\text{ pf}}$: just use that a convex fnl is lsc under weak convergence; so $F^\nu = \partial x^\nu/\partial X \to F^\infty = \partial x^\infty/\partial X$ and $\det F^\nu \to \det F^\infty$ wkly in $L^p \times L^{p/2}$ implies

$$\liminf_\nu \int_\Omega \varphi(F^\nu, \det F^\nu) \geq \int_\Omega \varphi(F^\infty, \det F^\infty)$$

when φ is convex (with appropriate growth).

It remains to explain the wk cont'y of $\det Dx$.
Main pt:

(*) <u>$\det Dx$ can be expressed as a divergence</u>

[How could we have guessed this? Well,

$$\int_\Omega \det Dx = \text{area of image of } \Omega \text{, if } \det Dx > 0.$$

So integral depends only on bdry data. How else could this happen, but by $\det Dx$ being a divergence? Note corollary to the preceding: EL eqn for var'l plm "min $\int \det Dx$" must be an identity. Therefore we say $\det Dx$ is a "null-Lagrangian".]

Pf of (*), via exterior calculus:

$$d(x_1 \wedge dx_2) = dx_1 \wedge dx_2 = (\det Dx)\, dX_1 \wedge dX_2 .$$

Same pf, in coordinates:

$$\det Dx = \partial_1 x_1 \partial_2 x_2 - \partial_1 x_2 \partial_2 x_1$$

$$= \partial_1 (x_1 \partial_2 x_2) - \partial_2 (x_1 \partial_1 x_2)$$

To see why (*) implies wk convergence of $\det Dx^\nu$ to $\det Dx^\infty$ when $x^\nu \rightharpoonup x^\infty$ wkly $W^{1,p}$ $(p>2)$, it's suff't to prove convergence as <u>distributions</u>. So let

φ be smooth with cpt supt. Then

$$\int_\Omega (\det Dx)\,\varphi = \int_\Omega -x_1\,\partial_2 x_2\,\partial_1\varphi + x_1\partial_1 x_2\partial_2\varphi$$

If $x^\nu \to x^\infty$ wkly $W^{1,p}$ then

$$\int_\Omega x_1^\nu\,\partial_2 x_2^\nu\,\partial_1\varphi \;\to\; \int_\Omega x_1^\infty\,\partial_2 x^\infty\,\partial_1\varphi$$

since $x_1^\nu \to x_1^\infty$ strongly in L^p (using Rellich)
$\partial_2 x_2^\nu \to \partial_2 x_2^\infty$ wkly in L^p (by hypothesis)

+ similarly for the other term.

Extnsn of above discn to 3D elasticity: each 2×2 minor is wkly cont's in $L^{p/2}$ under wk conv. in $W^{1,p}$ $(p>2)$; also the 3×3 det is wkly conts in $L^{p/3}$ under wk conv in $W^{1,p}$ $(p>3)$.

Suggested exercises:

(1) We showed that if $x^n \to x^\infty$ wkly in $W^{1,p}(\Omega)$, $p>2$ (for maps $\mathbb{R}^2\to\mathbb{R}^2$; by defn this means $\int_\Omega |Dx^n|^p < C$ with C indep of n, and $\int_\Omega \langle Dx^n,\varphi\rangle \to \int_\Omega \langle Dx^\infty,\varphi\rangle$ for any L^p vector field φ on Ω) then $\det Dx^n \to \det Dx^\infty$

wkly $L^{p/2}$. When $\{x^n\}$ are more singular strange things can happen due to "cavitation." Explore this by considering the maps $x^n : B_1 \to B_1$ (where B_1 = unit ball in $\mathbb{R}^2$) such that

$$x^n(X) = \begin{cases} \dfrac{1-\rho_n}{\rho_n} X & \text{for } |X| \le \rho_n \\ \left(\dfrac{1-2\rho_n}{1-\rho_n} + \dfrac{\rho_n}{1-\rho_n} |X| \right) \dfrac{X}{|X|} & \text{for } \rho_n \le |X| < 1 \end{cases}$$

so x^n "blows up" B_{ρ_n} to $B_{1-\rho_n}$ and "squashes" $B_1 \setminus B_{\rho_n}$ to $B_1 \setminus B_{1-\rho_n}$. (Here ρ_n can be any sequence converging to 0.)

a) Show that $x^n(X) \to \dfrac{X}{|X|}$ a.e.

b) Show that $\det Dx^n$ converges as a distribution to a point mass located at 0

c) For which p does $\displaystyle\int_{B_1} |Dx^n|^p \, dX$ stay bounded as $n \to \infty$?

[Comment: for additional examples, some similar and others rather different, see §7 of J Ball + F Murat, J Functional Analysis 58 (1984) 225-253. That section is independent of the rest of the paper.]

(2) Show that if $\int_\Omega g(\det Du)\,dx$ is lower semicontinuous (for $\Omega \subset \mathbb{R}^2$ + $u : \Omega \to \mathbb{R}^2$) then g must be convex.

<u>Hint</u>: it suffices to show that $g(\theta a + (1-\theta)b) \leq \theta g(a) + (1-\theta) g(b)$ for $0 < \theta < 1$ and $a, b \in \mathbb{R}$. Do this by considering $u_\varepsilon(x,y) = (u_{1\varepsilon}, u_{2\varepsilon})$ defined by

$$u_{1\varepsilon}(x,y) = (\theta a + (1-\theta) b)x + \varepsilon\varphi(x/\varepsilon)$$

$$u_{2\varepsilon}(x,y) = y$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is periodic with period 1 such that

$$\varphi(t) = \begin{cases} t(1-\theta)(a-b) & 0 < t < \theta \\ (1-t)\theta(a-b) & \theta < t < 1 \end{cases}.$$

(3) It can be nontrivial to determine whether a given function $W(F)$ is polyconvex or not. As an example, show that in the 2x2 setting

$$W(F) = \begin{cases} 1 + |F|^2 & \text{if } \rho(F) \geq 1 \\ 2\rho(F) - 2|\det F| & \text{if } \rho(F) \leq 1 \end{cases}$$

is polyconvex, when $|F|^2 = \sum_{i,j=1}^{2} |F_{ij}|^2$ and

$$\rho(F) = (|F|^2 + 2|\det F|)^{1/2}.$$

(This example arose as the "relaxation" of a nonconvex problem from "optimal design", in my work with Strang in the 80's.)

<u>Hint</u>: show that $W(F) = g(F, \det F)$ where

$$g(F,t) = \max_{\alpha=\pm 1} \left\{ f\left(\left[|F|^2 + 2\alpha \det F \right]^{1/2} \right) - 2\alpha t \right\}.$$

in which

$$f(t) = \begin{cases} 1+t^2 & t \geq 1 \\ 2t & t \leq 1 \end{cases},$$

Then check that $g(F,t)$ is a convex function of F and t.