Calculus of Variations, Lecture 2, 9/23/09

(There was no class on 9/16.)

Today's topic: convex duality. What is it?

a) Suppose we're interested in the min value of a convex optimization, eg

$$\inf_{\partial c} \int W(\nabla u)\, dx$$

where W is convex. Upper bds are easy (any choice of u gives one). But what about lower bounds? The convex dual provides a systematic approach.

b) Suppose we're interested in a convex but non-smooth variational problem like

$$\min_{\substack{\int_\Omega u\, dx = 1 \\ u = 0 \text{ at } \partial\Omega}} \int_\Omega |\nabla u|$$

The convex dual provides nec + suff't conds for optimality (playing a role analogous to the Euler-Lagrange eqn of a smoother convex var'l pbm)

Many key ideas are already visible in linear programming.

Consider (to fix ideas) the "primal problem"

$$(\mathcal{P}) \qquad \min \sum_{j=1}^{n} c_j x_j \qquad x \in \mathbb{R}^n$$
$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad 1 \leq i \leq m$$
$$x_j \geq 0$$

We can derive "trivial lower bd" on opt'l value by taking lin combns of constraints: if $y_i \geq 0$ and $\sum_{i=1}^{m} a_{ij} y_i \leq c_j$ then

$$y_i \sum_j a_{ij} x_j \geq b_i y_i \underset{\text{adding}}{\Rightarrow} \sum_{j=1}^{n} c_j x_j \geq \sum_{i=1}^{m} b_i y_i$$

The best "trivial lower bd" is obtained by maximization:

$$(\mathcal{D}) \qquad \max \sum_{i=1}^{m} b_i y_i$$
$$\sum_{i=1}^{m} a_{ij} y_i \leq c_j$$
$$y_i \geq 0$$

Duality thm of lin progr says

$$\max \mathcal{D} = \min \mathcal{P}$$

ie even the exact value of $\min \mathcal{P}$ can be achieved by a "trivial lower bd." (Proof of this thm is not trivial: see P. Lax's Lin Alg book for an attractive approach close to spirit of this lecture. But any lin progr. text will have a proof; I like the book by Chvatal.)

Note: if y^* solves dual + x^* solves primal then (by duality thm) $\sum' c_j x_j^* = \sum' b_i y_i^*$.
Examining the calcn on pg 2 we see that certain relns must hold:

$\forall i$: $y_i^* \geq 0$ + $\sum_j' a_{ij} x_j^* \geq b_i$ with equality in at least one of the two

$\forall j$ $x_j^* \geq 0$ + $\sum' a_{ij} y_i^* \leq c_j$ with equality in at least one of the two

Thus: existence of solution y^* to these "complementary slackness conditions" replaces the Euler-Lagrange eqns (nb for a smooth convex problem, any crit pt is a minimum).

Here's a basic pde example of two problems in duality. Suppose $f: \partial\Omega \to \mathbb{R}$ satisfies $\int_{\partial\Omega} f\, ds = 0$. Consider

$$(P) \qquad \min_u \int_\Omega \tfrac{1}{2}|\nabla u|^2 - \int_{\partial\Omega} uf\, ds$$

$$(\mathcal{D}) \qquad \max_{\substack{\operatorname{div}\sigma = 0 \\ \sigma\cdot n = f \text{ at } \partial\Omega}} \; -\frac{1}{2}\int_\Omega |\sigma|^2$$

[Rmk: we use here the fact that $\operatorname{div}\sigma = 0$, $\sigma\in L^2 \Rightarrow$ $\sigma\cdot n$ has a well-defined trace on bdry, for which Green's thm holds: $\int_{\partial\Omega}(\sigma\cdot n)u = \int_\Omega \langle\sigma,\nabla u\rangle + \cancel{\int_\Omega u \operatorname{div}\sigma}$

for all $u\in H^1(\Omega)$. In general $\sigma\cdot n \in H^{-1/2}(\partial\Omega)$ = dual to $H^{1/2}(\partial\Omega)$, since $H^{1/2}(\partial\Omega)$ = exact space of traces of H^1 fns. Special case $\Omega\subset\mathbb{R}^2$ is easiest since $\operatorname{div}\sigma = 0 \Rightarrow \sigma = (\nabla\varphi)^\perp$ and $\sigma\cdot n = \partial_{\tan}\varphi\big|_{\partial\Omega}$.]

The primal + dual have same relation as before:

① if $\left\{\begin{matrix}\sigma \text{ admissible} \\ \text{for } \mathcal{D}\end{matrix}\right\}$ + $\left\{\begin{matrix}u \text{ admissible} \\ \text{for } \mathcal{P}\end{matrix}\right\}$ then

value of $\mathcal{D}$ at σ $\leq$ value of $\mathcal{P}$ at u,

② and equality holds when u solves $\mathcal{P}$ + σ solves $\mathcal{D}$.

Proof in this case is elementary: to see ①, expand

$$\int_\Omega \frac{1}{2}|\sigma-\nabla u|^2 \geq 0$$

to see $\int \frac{1}{2}|\sigma|^2 + \frac{1}{2}|\nabla u|^2 - \langle \sigma, \nabla u\rangle \geq 0$.

Now use $\operatorname{div}\sigma = 0$, $\sigma.n = f$ to get

$$-\frac{1}{2}\int_\Omega |\sigma|^2 \leq \int_\Omega \frac{1}{2}|\nabla u|^2 - \int_{\partial\Omega} u f$$

To see ②, observe that if u^* solves primal + $\sigma^* = \nabla u^*$ then preceding inequalities are $=$.

Thus: in this case the cond. for equality

$$\begin{matrix} \operatorname{div}\sigma = 0 \\ \sigma . n = 0 \end{matrix} \quad + \quad \sigma = \nabla u$$

are just a rewrite of the EL eqn for $\mathcal{P}$

Rmk: When doing numerical calculations by finite element method it can be difficult to know how good an approxn you have obtained. A "primal-dual" method can study $\mathcal{P}$ + $\mathcal{D}$ simultaneously. Lemma: if $\hat{\sigma}$ is admissible for $\mathcal{D}$ + $\hat{u}$ is admissible for $\mathcal{P}$ and

$$(\text{value of } \mathcal{P} \text{ at } \hat{u}) - (\text{value of } \mathcal{D} \text{ at } \hat{\sigma}) < \delta$$

then $\frac{1}{2}\int |\nabla\hat{u} - \nabla u^*|^2 \leq \delta$ and $\frac{1}{2}\int |\hat{\sigma} - \sigma^*|^2 \leq \delta$

where u^* and $\sigma^* = \nabla u^*$ are the solns of $\mathcal{P}$ + $\mathcal{D}$.

(Proof: exercise.)

How could we have found the dual problem systematically? Ans: dual pairs are assoc to "saddle pt" var'l pbms (ie to switching a max + min.) Explain by considering the more general problem

$$(\mathcal{P}) \quad \min_u \int_\Omega W(\nabla u) - \int_{\partial\Omega} u \cdot f$$

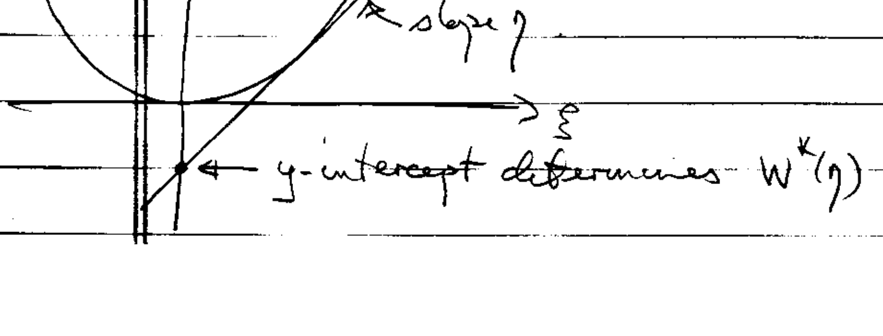
with $W(\xi)$ convex.

Key pt: W convex $\Leftrightarrow$ its graph is envelope of supporting hyperplanes

$$\Leftrightarrow \quad W(\xi) = \sup_\eta \langle \eta, \xi \rangle - W^*(\eta)$$

where W^* (the "Fenchel transform" of W) is given by

$$W^*(\eta) = \sup_\xi \langle \xi, \eta \rangle - W(\xi).$$



So:
$$\min_u \int_\Omega W(\nabla u) - \int_{\partial\Omega} u\cdot f$$
$$= \min_u \max_{\sigma(x)} \int_\Omega \langle \sigma, \nabla u\rangle - W^*(\sigma) - \int_{\partial\Omega} u\cdot f$$
$$= \min_u \max_{\sigma(x)} \int_{\partial\Omega} (\sigma\cdot n)u - f\cdot u + \int_\Omega -(\operatorname{div}\sigma)u - W^*(\sigma)$$

Claim: we can switch min + max. (Rtn to this soon).
$$= \max_{\sigma(x)} \min_u \int_{\partial\Omega} (\sigma\cdot n - f)u - \int_\Omega (\operatorname{div}\sigma)u + W^*(\sigma)$$
$$= \max_{\substack{\operatorname{div}\sigma = 0 \\ \sigma\cdot n = f}} -\int_\Omega W^*(\sigma)$$

since if $\operatorname{div}\sigma \neq 0$ or $\sigma\cdot n \neq f$ then min over u would be $-\infty$.

Why is min max = max min? As usual, an inequality is trivial

1st viewpt: $\min_y F(x,y) \leq F(x,y_0)$
$$\Rightarrow \max_x \min_y F(x,y) \leq \max_x F(x,y_0)$$
$$\Rightarrow \max_x \min_y F(x,y) \leq \min_y \max_x F(x,y)$$

2nd viewpt: suppose $\operatorname{div} \sigma = 0$, $\sigma \cdot n = f$. Integrate

$$W(\nabla u) \geq \langle \nabla u, \sigma \rangle - W^*(\sigma)$$

to get $$\int_\Omega W(\nabla u) - \int_{\partial\Omega} u \cdot f \geq -\int_\Omega W^*(\sigma).$$

The fact that we get equality (not inequality) is nontrivial in many cases. But if $\mathcal{P}$ or $\mathcal{D}$ has a clearly defined EL eqn then it will give us a direct pf. In present setting: if W is smooth enough that

$$\operatorname{div}\left(\frac{\partial W}{\partial \nabla u}\right) = 0 \qquad \frac{\partial W}{\partial \nabla u} \cdot n = f$$

has a soln, then we can take $\sigma = \dfrac{\partial W}{\partial \nabla u}$.

(Note: when $W(\xi) = \frac{1}{2}|\xi|^2$, $W^*(\sigma) = \frac{1}{2}|\sigma|^2$, so this example reduces to our quadratic one.)

Plan for rest of lecture:

a) give a more subtle example that still involves linear pde

b) discuss some examples involving L^1-L^∞ duality (where a conventional Euler-Lagrange eqn can't be written)

Here's example (a): Let λ_0 be 1st Dirichlet eigenvalue of domain Ω:

$$\lambda_0 = \min_{u=0 \text{ at } \partial\Omega} \frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2} = \min_{\substack{u=0 \text{ at } \partial\Omega \\ \int_\Omega u^2 = 1}} \int_\Omega |\nabla u|^2 .$$

Upper bds are easy (consider any u). How about a scheme for proving lower bounds?

Step 1 Suff't to consider $u \geq 0$, since replacing u by $|u|$ leaves both $\int |\nabla u|^2$ + $\int u^2$ invariant. (Exercise)

Step 2 Let $\rho = u^2$ (ie let $u = \sqrt{\rho}$) + write defn of λ_0 in terms of ρ:

$$\lambda_0 = \min_{\substack{\int \rho \, dx = 1 \\ \rho \geq 0 \text{ in } \Omega \\ \rho = 0 \text{ at } \partial\Omega}} \int_\Omega \frac{|\nabla \rho|^2}{4\rho} \, dx .$$

step 3 Observation: the fn $|\xi|^2/4t$ is a convex fn of (ξ, t). In fact

$$\xi^2/4t = \max_{\sigma} \langle \sigma, \xi \rangle - t|\sigma|^2$$

So
$$\lambda_0 = \min_{\substack{\int \rho = 1 \\ \rho \geq 0 \\ \rho = 0 \text{ at } \partial\Omega}} \max_{\sigma(x)} \int_\Omega \langle \sigma, \nabla\rho \rangle - \rho|\sigma|^2.$$

step 4 Proceed as usual: switch min + max, and use min over ρ to get constraints on σ

$$\lambda_0 = \max_{\sigma} \min_{\substack{\int \rho = 1 \\ \rho \geq 0 \\ \rho = 0 \text{ at } \partial\Omega}} \cancel{\int_{\partial\Omega} \rho \sigma \cdot n} - \int (\operatorname{div}\sigma + |\sigma|^2)\rho$$

$$= \max_{\operatorname{div}\sigma + |\sigma|^2 = -\mu \text{ constant}} \mu$$

= largest constant μ st vector field σ st $\operatorname{div}\sigma + |\sigma|^2 = -\mu$.

step 5 Is the max/min right? Sure!
$\max_\sigma \langle \sigma, \xi \rangle - t|\sigma|^2$ is achieved when $\xi = 2t\sigma$.

So best σ is $\frac{1}{2\rho}\nabla\rho$ where $\rho = u^2$ + u is 1st Dirichlet eigenfn. Direct calcn $\Rightarrow$ it is admissible for dual pbm + achieves its opt'l value. (Exercise.)

Here's <u>example (b)</u>: What can we say about

$$(*) \qquad \min_{\sigma} \left\{ \|\sigma\|_\infty \text{ st } \operatorname{div}\sigma = 1 \text{ on } \Omega \right\}.$$

where (to fix ideas) Ω is a bdd domain in $\mathbb{R}^2$. (Interpretation: it's raining uniformly on Ω. How can rain flow to bdry with least possible local accumulation?)

1st observation: it's equivalent to solve

$$(**) \qquad \min \lambda$$
$$|\sigma| \le 1$$
$$\operatorname{div}\sigma = \lambda \text{ (constant)}$$

since if λ_{max} is opt'l value of (**) then

$$\begin{matrix} \operatorname{div}\sigma = \lambda \text{ (const)} \\ |\sigma| \le 1 \end{matrix} \quad \Rightarrow \quad \lambda \le \lambda_{max}$$

so $\quad \operatorname{div}\left(\frac{\sigma}{\lambda}\right) = 1 \quad \Rightarrow \quad \frac{1}{\lambda} \geq \frac{1}{\lambda_{max}}$

$\left|\frac{\sigma}{\lambda}\right| \leq \frac{1}{\lambda}$

so $1/\lambda_{max}$ is opt'l value for (*).

Identify dual of (**) by usual argument:

$$\max_{|\sigma| \leq 1} \min_{\substack{u=0 \text{ at } \partial\Omega \\ \int u = 1}} \int_\Omega \langle \nabla u, \sigma \rangle = \min_{\substack{u=0 \text{ at } \partial\Omega \\ \int u = 1}} \max_{|\sigma| \leq 1} \int_\Omega \langle \sigma, \nabla u \rangle$$

min over $u = -\infty$
unless $\operatorname{div} \sigma = \lambda$ constant

$\Downarrow$

$$\max_{\substack{|\sigma| \leq 1 \\ \operatorname{div}\sigma = \lambda \text{ constant}}} -\lambda$$

opt'l σ has
$\langle \sigma, \nabla u \rangle = |\nabla u|$

$\Downarrow$

$$\min_{\substack{u=0 \text{ at } \partial\Omega \\ \int_\Omega u = 1}} \int_\Omega |\nabla u| .$$

Is max min = min max? Inequality is elementary as usual

$$\left.\begin{array}{l} |\sigma| \leq 1 \\ u = 0 \text{ at } \partial\Omega \\ \operatorname{div}\sigma = \lambda \\ \int u = 1 \end{array}\right\} \Rightarrow \int \langle \sigma, \nabla u \rangle \leq \int |\nabla u|$$

$$\int \langle \sigma, \nabla u \rangle = -\lambda$$

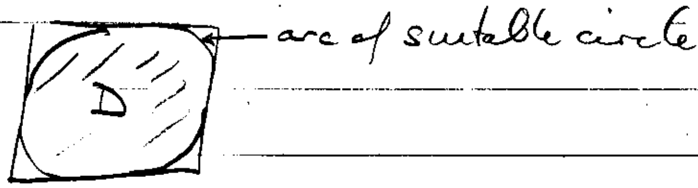
However equality is not simple to prove. Moreover opt'l u is rather singular – it's the characteristic fn of a set (see below). That max min = min max can be proved using results from book by Ekeland + Temam (Convex Analysis + Variational Problems). Similar L^1-L^∞ duality plms arise in plasticity & other areas of mechanics; see eg book by Duvaut + Lions (Inequalities in Mechanics + Physics).

Interesting feature of L^1-L^∞ pairs: one is typically much easier to solve than the other. Here we have

$$(***) \qquad \min_{\substack{u=0 \text{ at } \partial\Omega \\ \int_\Omega u = 1}} \int_\Omega |\nabla u| = \min_{D\subset\Omega} \frac{\text{length}(\partial D)}{\text{Area}(D)}$$

= soln of a geometry plm!

Example for Ω = square:



(See G. Strang, "A minimax problem in plasticity theory", Springer Lecture Notes in Math 701, 1979, 319–333)

Sketch proof of (***). A key ingredient is the

<u>coarea formula</u> $\quad \int f(x)\,|\nabla u|\,dx = \int \left(\int_{u=t} f\,ds \right) dt$

(easy to justify if u is nice enough - more or less this is the "method of shells" from Calc III).

<u>Step 1</u> LHS of (***) $= \min\limits_{u=0 \text{ at } \partial\Omega} \dfrac{\int_\Omega |\nabla u|}{\int_\Omega u}$

(easy)

<u>Step 2</u> May assume $u \geq 0$ (since replacing $u \to |u|$ leaves $\int |\nabla u|$ invariant + increases $\int u$)

<u>Step 3</u> For $u \geq 0$, $\int u\,dx = \int_0^\infty \text{Area}\{u \geq t\}\,dt$

since

$$\int_\Omega u\,dx = \int_\Omega \int_0^{u(x)} 1\,dt\,dx$$

(now use Fubini's Thm). On other hand

$$\int |\nabla u|\,dx = \int_0^\infty \text{length}\{u = t\}\,dt$$

(coarea formula with $f=1$). So: if RHS of (**) has min α then

$$\text{length}\{u=t\} \geq \alpha \text{ Area}\{u \geq t\} \quad \underline{\text{for all } t}$$

$$\Rightarrow \quad \int_\Omega |\nabla u| \geq \alpha \int_\Omega u$$

$$\Rightarrow \quad \text{LHS of } (***) \geq \alpha.$$

For more discn of closely related pblms see recent paper by G. Strang, "Maximum flows + minimum cuts in the plane" (J. Global Optzn., in press, There's a preprint on Strang's website.) I'll take just one item from there: a very efficient pf (due to Grieser) of

<u>Cheeger's inequality</u>: if $\lambda_0 = 1^{st}$ Dirichlet eigenvalue

$$h = \min_{D \subset \Omega} \frac{\text{length}(\partial D)}{\text{Area}(D)}$$

Then $\frac{h^2}{4} \leq \lambda_0$.

<u>Pf</u>: From prior discn, $\exists\ \sigma$ st $|\sigma| \leq 1$ + $\operatorname{div} \sigma = h$. Let u_0 be 1^{st} dir eigenfn. Then

$$h \int_\Omega u_0^2 = \int_\Omega \operatorname{div}\sigma \cdot u_0^2 = -\int_\Omega 2u_0 \langle \sigma, \nabla u_0 \rangle$$

$$\leq 2 \int |u_0| \, |\nabla u_0|$$

$$\leq 2 \left(\int u_0^2 \right)^{1/2} \left(\int |\nabla u_0|^2 \right)^{1/2}$$

$$\Rightarrow \quad h \leq 2 \frac{\left(\int |\nabla u_0|^2 \right)^{1/2}}{\left(\int u_0^2 \right)^{1/2}} = 2 \lambda_0^{1/2}.$$

Some exercises:

1) Show that the convex dual of

$$\min_{u = u_0 \text{ at } \partial\Omega} \int \tfrac{1}{2} |\nabla u|^2$$

is $$\max_{\operatorname{div}\sigma = 0} \int_{\partial\Omega} (\sigma \cdot n) u_0 - \tfrac{1}{2} \int_\Omega |\sigma|^2.$$

2) Show that the var'l problems

(P) $$\min_{\substack{\operatorname{div}\sigma = F \text{ in } \Omega \\ \sigma \cdot n = f \text{ at } \partial\Omega}} \int_\Omega |\sigma|$$

$$(\mathcal{D}) \quad \max_{|\nabla u|_\infty \le 1} \int_{\partial\Omega} u \cdot f \, ds - \int_\Omega u \cdot F \, dx$$

are a dual pair, if $\int_\Omega F\,dx = \int_{\partial\Omega} f$.

How should σ + ∇u be related if equality is to hold?

(Hint: $\max_{\sigma\in\mathbb{R}^n} \langle \xi, \sigma\rangle - |\sigma| = \begin{cases} 0 & \text{if } |\xi| \le 1 \\ +\infty & \text{if } |\xi| > 1. \end{cases}$)

Rmk: if $\Omega \subset \mathbb{R}^2$ + $F=0$ then $\mathcal{P}$ can be solved explicitly in simple cases using the co-area formula. Why?

3) When we study homogenization we'll consider a periodic "conductivity" $a(x)$, + we'll learn that the assoc "effective conductivity" a_{eff} satisfies

$$\langle a_{eff}\,\xi, \xi\rangle = \min_{\varphi \text{ periodic}} ⨍ \langle a(x)(\xi + \nabla\varphi), \xi + \nabla\varphi\rangle$$

(a_{eff} is in general a matrix; ⨍ denotes the average of a periodic function). Show using duality that

$$\langle a_{eff}\,\xi,\xi\rangle = \max_{\substack{\mathrm{div}\,\sigma=0\\ \sigma\ \text{periodic}}} \fint \langle 2\sigma,\xi\rangle - \langle a^{-1}(x)\sigma,\sigma\rangle$$

This can alternatively be written as

$$2\langle\bar\sigma,\xi\rangle - \langle a_{eff}\,\xi,\xi\rangle = \min_{\substack{\mathrm{div}\,\sigma=0\\ \sigma\ \text{periodic}\\ \fint\sigma=\bar\sigma}} \fint\langle a^{-1}(x)\sigma,\sigma\rangle$$

for any $\xi,\bar\sigma\in\mathbb{R}^n$. Optimizing over ξ, we get

$$\langle a_{eff}^{-1}\,\eta,\eta\rangle = \min_{\substack{\mathrm{div}\,\sigma=0\\ \sigma\ \text{periodic}\\ \fint\sigma=\eta}} \fint\langle a^{-1}(x)\sigma,\sigma\rangle$$