Calculus of Variations - Lecture 2, 1/30/2017

Today: start convex duality (we'll spend 2 or 3 lectures on this + related material).

Rough plan:

- two distinct goals...
- brief discn of linear programming
- a basic pde example of convex duality
- The Fenchel transform, and derivation of dual problem by min max → max min
- some specific, interesting examples of convex duality in pde settings
- The "calibration method" (not convex duality exactly, but closely related)

Throughout our discussion, 2 goals will be intertwined:

(1) We're often interested in the min value of a convex optzn, eg

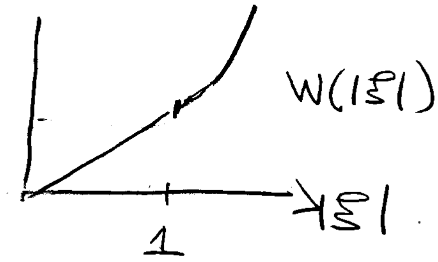
$$\inf_{bc} \int_\Omega W(\nabla u)\, dx$$

with W convex. Upper bounds are easy

(any choice of u gives one). But what about lower bounds? The convex dual provides a systematic approach.

(2) We're sometimes interested in non-smooth var'l pbms, whose EL doesn't quite make sense; for example

$$\min_{bc} \int_\Omega W(\nabla u)\,dx \quad \text{with } W(\xi) = \begin{cases} 2|\xi| & |\xi| \le 1 \\ 1+|\xi|^2 & |\xi| \ge 1 \end{cases}$$



or $\displaystyle \min_{\substack{\int_\Omega u\,dx = 1 \\ u=0 \text{ at } \partial\Omega}} \int_\Omega |\nabla u|$ (to be discussed in detail next week).

In such cases, EL eqn says, formally

$$\operatorname{div}\left(\frac{\partial W}{\partial \xi}(\nabla u)\right) = 0 \quad \rightarrow \quad \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0$$

in 2nd example, or in 1st example if $|\nabla u| < 1$.

<u>But this doesn't make sense if $\nabla u = 0$. Convex duality provides a substitute for EL eqn in non-smooth setting.</u>

Many key ideas are already visible in <u>linear programming</u>. Consider (to fix ideas) the "primal problem"

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

We can derive a "trivial lower bound" on the optimal value by taking a linear combination of the constraints: if $y_i \geq 0$ and $\sum_{i=1}^{m} a_{ij} y_i \leq c_j$ then

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

The best "trivial lower bd" is obtained by optimizing this result:

$$(\mathcal{D}) \quad \max_{\substack{\sum_{i=1}^{m} a_{ij} y_i \le c_j \\ y_i \ge 0}} \quad \sum_{i=1}^{m} b_i y_i$$

The <u>duality theorem</u> of linear programming says

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

i.e the <u>exact value</u> $\min \mathcal{P}$ is achieved by a well-chosen "trivial lower bd".
(The proof is <u>not</u> trivial. See P. Lax's Linear Alg book for a nice proof that's in the spirit of my discussion. But <u>any</u> linear programming text will give a proof. I like the one by Chvatal.)

<u>Note</u>: if y^* solves $\mathcal{D}$ and x^* solves $\mathcal{P}$ then (by duality thm) $\sum_j c_j x_j^* = \sum_i b_i y_i^*$.

Examining prev calcn we see that

$\forall i$, $y_i^* \ge 0$ and $\sum_j a_{ij} x_j^* \ge b_i$ with equality in at least one of these

and

$\forall j$, $x_j^* \geq 0$ and $\sum_i a_{ij} y_i^* \leq c_j$ with equality in at least one of these

These "complementary slackness" condns play the role in linear programming that the EL eqn plays in smooth + convex pde-type problems.

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

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

$$(\mathcal{D}) \quad \max_{\substack{\operatorname{div} \sigma = 0 \text{ in } \Omega \\ \sigma \cdot n = f \text{ at } \partial\Omega}} -\tfrac{1}{2} \int_\Omega |\sigma|^2 \, dx$$

These problems have the same relationship to each other as our $\mathcal{P}$ + $\mathcal{D}$ in linear programming:

(a) if σ is admissible for $\mathcal{D}$ and u is admissible for $\mathcal{P}$ then:

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

(b) equality holds when u^* solves $\mathcal{P}$ and σ^* solves $\mathcal{D}$

The following proof is elementary + quick (but maybe it doesn't generalize so easily; we'll give a more general treatment soon).

To see (a), expand

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

to see

$$\int_\Omega \tfrac{1}{2}|\sigma|^2 + \tfrac{1}{2}|\nabla u|^2 - \langle \sigma, \nabla u \rangle \, dx \geq 0.$$

Then use $\operatorname{div} \sigma = 0$ in Ω, $\sigma \cdot n = f$ at $\partial\Omega$ to get

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

To see (b), observe that if u^* solves

(P) and $\sigma^* = \nabla u^*$ then (from pf of (a)) σ^* solves (D) [since all the ineqs are equalities in the discn of (a)]. Note: this σ^* is the only soln of D, since the fnl $\int |\sigma|^2$ is strictly convex.

Note: in this case the chargn of the optimal choices

$$\operatorname{div} \sigma^* = 0 \text{ in } \Omega \qquad \sigma^* = \nabla u^*$$
$$\sigma^* \cdot n = f \text{ at } \partial\Omega$$

amounts to a rewrite of the EL eqn for P ($\Delta u^* = 0$ in Ω, $\partial u^*/\partial n = f$ at $\partial\Omega$). [When the EL of P makes sense, the complementary slackness condn is always equivalent to the EL eqn.]

Digression 1: I skipped over the question: what is the proper function space for D? [Everything is fine if σ is smooth – but how to know there's an optimal σ that's smooth?]. Given form of D, natural space to look (eg using the Direct Method) is

$$X = \{\sigma : \int_\Omega |\sigma|^2 < \infty \text{ and } \operatorname{div} \sigma = 0 \text{ in } \Omega\}.$$

Does the constraint "$\sigma \cdot n = f$ at $\partial\Omega$" make sense? Answer is yes: there's a cont's map

$$X \longrightarrow H^{-1/2}(\partial\Omega).$$

taking $\sigma \to \sigma \cdot n\big|_{\partial\Omega}$ when σ is smooth

and Green's formula holds in the sense that

$$\int_{\partial\Omega} (\sigma \cdot n)\, u = \int_{\Omega} \langle \sigma, \nabla u \rangle + \cancel{\int_{\Omega} u \operatorname{div} \sigma} \quad \text{for all } u \in H^1(\Omega)$$

Here $H^{-1/2}(\partial\Omega)$ = dual of $H^{1/2}(\partial\Omega)$ in L^2 inner product

$H^{1/2}(\partial\Omega)$ = exact space of boundary traces of $H^1(\Omega)$ functions.

The special case $\Omega \subset \mathbb{R}^2$ is more elementary since $\operatorname{div} \sigma = 0 \Rightarrow \sigma = (\nabla \varphi)^{\perp}$, and $\sigma \cdot n\big|_{\partial\Omega} = \partial_{\tan} \varphi\big|_{\partial\Omega}$.

<u>Digression 2</u>: When doing numerical calculations by finite element method it can be difficult to know how good an approx'n of the sol'n you have obtained. A "primal-dual" method studies $\mathcal{P}$ and $\mathcal{D}$ simultaneously. Advantage of this is evident from:

<u>Lemma</u>: If $\hat{\sigma}$ is admissible for $\mathcal{D}$ and $\hat{u}$ is admissible for $\mathcal{P}$ and

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

Then

$$\frac{1}{2}\int_\Omega |\nabla\hat{u} - \nabla u^*|^2 \leq \delta \quad \text{and} \quad \frac{1}{2}\int_\Omega |\hat{\sigma} - \sigma^*|^2 \leq \delta$$

where u^*, σ^* are the solns of $\mathcal{P}$ and $\mathcal{D}$. (Proof: exercise. Note the similarity to Exercise 6 at the end of Lecture 1.)

<u>How to find dual problems systematically?</u>

Central idea: a convex optimization can be expressed as a min max. Switching the min + the max gives the dual pbm. (There can be more than one min/max repn of a given var'l pbm; different choices may lead to slightly different "dual" pbms.)

More detail now, focusing (to keep things simple) on

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

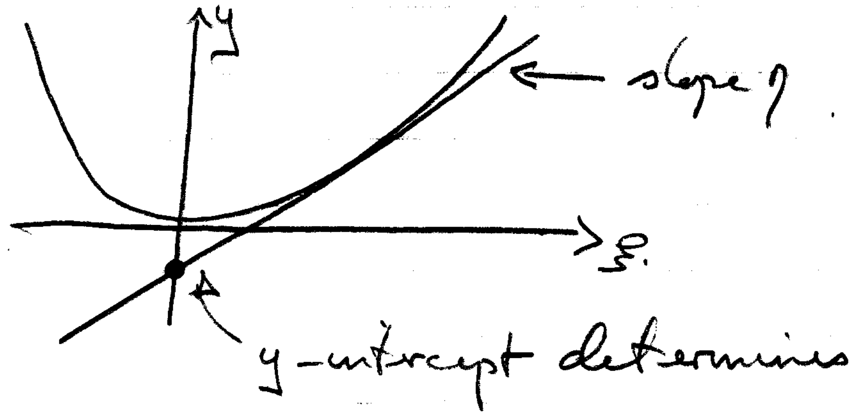
with $W(\xi)$ convex and $\int f \, ds = 0$.

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

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

where $W^*(\eta)$ (the "Fenchel transform" of W) is defined by

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



So:
$$\min_u \int_\Omega W(\nabla u)\, dx - \int_{\partial\Omega} u \cdot f \, ds$$
$$= \min_u \max_\sigma \left\{ \int_\Omega \langle \sigma, \nabla u \rangle - W^*(\sigma)\, dx - \int_{\partial\Omega} u f \, ds. \right\}$$

<u>Claim</u>: we can switch min + max. (Return to this

soon). Then above becomes

$$\max_{\sigma} \min_{u} \int_{\partial\Omega} [(\sigma \cdot n) - f]\, u \, ds - \int_{\Omega} (\mathrm{div}\,\sigma)\, u + W^*(\sigma)$$

$$(D) \qquad = \max_{\substack{\mathrm{div}\,\sigma = 0 \text{ in } \Omega \\ \sigma\cdot n = f \text{ at } \partial\Omega}} - \int_{\Omega} W^*(\sigma)\, dx .$$

(we integrated by parts in the 1st line above, & observed that the min wrt u is $-\infty$ unless $\mathrm{div}\,\sigma = 0$ + $\sigma \cdot n = f$ to get 2nd line).

Why should min max = max min? An inequality is trivial, with no structural hypothesis:

1st viewpt: $\min_y F(x,y) \le F(x, y_0)$.

$$\Rightarrow \max_x \min_y F(x,y) \le \max_x F(x,y)$$

$$\Rightarrow \max_x \min_y F(x,y) \le \min_y \max_x F(x,y).$$

2nd viewpt: if $\mathrm{div}\,\sigma = 0$ and $\sigma \cdot n = f$ then integr of the ptwise inequality

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

gives

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

Either way, we see that for every σ (admissible for $\mathcal{D}$) we get a lower bd for $\mathcal{P}$.

The fact that we get equality, ie $\max \mathcal{D} = \min \mathcal{P}$, is nontrivial in general. Viewed as a saddle pt principle ($\min_x \max_y F(x,y) = \max_y \min_x F(x,y)$) it requires some conds on F (typically: concave in y, convex in X, and a little more - see eg book by Ekeland + Temam).

But: if $\mathcal{P}$ has a sensible EL eqn then we can use it to give a direct proof. In present setting: suppose W is convex + smooth with p^{th} power growth at ∞ ($W(\xi) \sim |\xi|^p$ as $|\xi| \to \infty$ and $\partial W / \partial \xi \sim |\xi|^{p-1}$ as $|\xi| \to \infty$). Then minimizer is a weak soln of EL eqn

$$\mathrm{div}\left(\frac{\partial W}{\partial \nabla u}\right) = 0 \text{ in } \Omega, \quad \frac{\partial W}{\partial \nabla u} \cdot n = f \text{ at } \partial\Omega.$$

Soln of dual pbm is then $\sigma^* = \left.\dfrac{\partial W}{\partial \nabla u}\right|_{u=u^*}$.

Since substn of this into $\mathcal{D}$ gives $\inf \mathcal{P}$, we get an explicit pf that $\inf \mathcal{P} = \sup \mathcal{D}$ without need for any min-max Thm. (Why is value of $\mathcal{D}$ at this σ^* equal to $\inf \mathcal{P}$? Exercise.)

Note: When $W(\xi) = \frac{1}{2}|\xi|^2$, $W^*(\sigma) = \frac{1}{2}|\sigma|^2$ and this discussion reduces to our prior quadratic one.

Digression: In discussing the "Direct Method" we used a theorem from fnl analysis to show that

$$E[u] = \int_\Omega W(\nabla u)\, dx.$$

is lower semicontinuous under wk conv of u (if W is convex with pth power growth, $1<p<\infty$). The Fenchel transform gives a different, more intuitive proof of lower semicontinuity. In fact

$$E[u] = \sup_{\sigma(x)} \int_\Omega \langle \sigma, \nabla u\rangle - W^*(\sigma)\, dx$$

and if σ is fixed then

$$u \longrightarrow \int_\Omega \langle \sigma, \nabla u\rangle - W^*(\sigma)\, dx$$

is cont's under weak convergence (in $W^{1,p}$) since it is <u>linear</u> in ∇u. So

$E[u]$ = max of fnls that are cont's (under wk convergence)

is lower <u>semicont's</u> under wk convergence.

Here's an example that's less standard, and therefore perhaps more interesting.

Let λ_0 = 1st Dirichlet eigenvalue of a domain $\Omega \subset \mathbb{R}^n$

$$= \min_{u=0 \text{ at } \partial\Omega} \frac{\int_\Omega |\nabla u|^2 dx}{\int_\Omega u^2 dx}$$

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

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

Step 1: Suff't to consider $u \geq 0$, since replacing u by $|u|$ leaves both $\int_\Omega |\nabla u|^2$ and $\int_\Omega u^2$ unchanged. (Exercise.)

Step 2: Let $\rho = u^2$ (i.e. let $u = \sqrt{\rho}$) and write the defn of λ_0 in terms of ρ:

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

Step 3: Observe that $\xi^2/4t$ is a convex fn of (ξ, t); in fact

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

So

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

<u>Step 4</u> Switch min + max to get a dual pblm:

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

$$= \max_{-[\operatorname{div}\sigma+|\sigma|^2]\geq\mu} \mu \quad \text{(constant!)}$$

= largest constant μ st $\exists$ vector field σ on Ω with $\operatorname{div}\sigma + |\sigma|^2 \leq -\mu$ ptwise.

<u>Step 5:</u> Is the max/min right? Sure! Observe that $\max_{\sigma} \langle\sigma,\xi\rangle - t|\sigma|^2$ is achieved when $\xi = 2t\sigma$,

So best σ is $\frac{1}{2\rho}\nabla\rho$ where $\rho = u^2$ and u is 1st Dirichlet eigenfn. Direct calculation $\Rightarrow$ this σ is admissible for proposed dual pblm and achieves its optimal value (Exercise!).

<u>Suggested exercises</u>

(1) [This pblm could have been at the end of

Lecture 1.] In Lecture 1 we used the Direct Method of the Calc of Varns to show that if $W(\xi)$ is convex with

$$C_1' + C_2'|\xi|^p \leq W(\xi) \leq C_1 + C_2|\xi|^p$$

then $\exists\ u_* \in W^{1,p}(\Omega)$ st u_* achieves

$$\min_{u=\varphi \text{ at } \partial\Omega} \int_\Omega W(\nabla u) + fu \; dx$$

Show that if W is differentiable and

$$\left|\frac{\partial W}{\partial \xi}\right| \leq C|\xi|^{p-1} \quad \text{as } |\xi| \to \infty$$

then the minimizer "satisfies the EL eqn" in the sense that

$$\int_\Omega \left\langle \frac{\partial W}{\partial \nabla u}(\nabla u_*), \nabla v \right\rangle + fv \; dx = 0$$

for any $v \in W^{1,p}(\Omega)$ st $v=0$ at $\partial\Omega$. (The main task is to justify putting the $\frac{d}{dt}$ under the integral in

$$\frac{d}{dt}\int_\Omega W(\nabla u + t\nabla v) + f(u+tv) \; dx \quad .)$$

(2) On pg 2.2 I mentioned, as non-smooth example, $\min_{u=\varphi \text{ at } \partial\Omega} \int_\Omega W(\nabla u)\,dx$ with

$$W(\nabla u) = \begin{cases} 2|\nabla u| & |\nabla u| \le 1 \\ 1+|\nabla u|^2 & |\nabla u| \geq 1. \end{cases}$$

What is the convex dual to this variational problem?

(3) Show that the dual of

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

is

$$\max_{\sigma} \int_{\partial\Omega} (\sigma\cdot n)\varphi - \tfrac{1}{2}\int_\Omega |\sigma|^2$$

(4) Do the exercise suggested on pg 2.9

(5) Do the exercise suggested near the top of pg 2.13