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

Rough plan:
- two distinct goals...
- brief discussion of linear programming
- a basic, ped example of convex duality
- the Fenchel duality form and derivation of dual problem by min max -> max min
- some specific, interesting examples of convex duality in real 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 opt prob, e.g.

$$\inf \int W(x) \, dx$$

where $W$ 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 variational problems, whose EL doesn't quite make sense for example

\[
\min_{bc \leq \frac{1}{5}} \int W(\gamma) \, dx \quad \text{with} \quad W(\gamma) = \begin{cases} 2.5, & |\gamma| < 1 \\ 1 + |\gamma|^2, & 1 \leq |\gamma| \end{cases}
\]

or

\[
\min \int |\gamma| \, dx \quad \text{with} \quad \int_{\Sigma} \gamma = 1, \quad \gamma = 0 \text{ a.e. on } \Sigma
\]

In such cases, EL epn says, formally

\[
\frac{\partial W}{\partial \gamma} (\gamma u) = 0 \quad \Rightarrow \quad \frac{\partial W}{\partial \gamma} \left( \frac{\gamma u}{17u} \right) = 0
\]

in 2nd example, or in 1st example if \( |\gamma| < 1 \).
But this doesn't make sense if \( \mathbf{u} = \mathbf{0} \).
Convex duality provides a substitute for EL in non-smooth settings.

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

\[
\begin{align*}
\min & \quad \sum_{i=1}^{n} c_i x_i \\
\text{subject to} & \quad \sum_{j=1}^{m} a_{ij} x_j = b_i, \quad 1 \leq i \leq m \\
& \quad x_j \geq 0
\end{align*}
\]

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

\[
\begin{align*}
\sum_{i=1}^{n} c_i x_i \geq \sum_{i=1}^{m} b_i y_i \Rightarrow \sum_{i=1}^{n} c_i x_i \geq \sum_{i=1}^{m} b_i y_i
\end{align*}
\]

The best "trivial lower bound" is obtained by optimizing this result.
The duality theorem of linear programming says

\[
\max \sum_{i=1}^{m} b_i y_i \\
\sum_{i=1}^{m} a_{ij} y_i \leq c_j \\
y_i \geq 0
\]

is the exact value \( \min P \) is achieved by a well-chosen "trivial" linear \( \text{rel.} \) \( \text{ld.} \). (The proof is not trivial. See P. Lax's Linear Algebra book for a nice proof that's in the spirit of my dissertation. But any linear programming text will give a proof. I like the one by Chvatal.)

Note: if \( y^* \) solves \( \max \) and \( x^* \) solves \( \min \)

then (by duality theorem) \( \sum_{i} c_i x^*_i = \sum_{i} b_i y^*_i \).

Examining prev calc we see that
\[ y_{i^*}^* \geq 0 \text{ and } \sum_{i} a_{ij} x^*_i = b_i \text{ with equality in at least one of these } \]
\[ x^* \geq 0 \text{ and } \sum_{i,j} a_{ij} x^* \leq c_i \text{ with equality in at least one of these.} \]

These "complementary slackness" conditions play the role in linear programming that the EL cone plays in smooth convex poly-type problems.

Here is a basic p.d.e. example of two problems in duality. Suppose \( f : \mathbb{R}^n \to \mathbb{R} \) satisfies \( \int f \, ds = b \). Consider

\[ \begin{align*}
\text{(b)} & \quad \text{min} \quad \int_{\mathbb{R}^2} \frac{1}{2} |x|^2 \, dx - \int_{\mathbb{R}^2} u \, ds \\
& \quad \text{subject to} \quad u_{x_1} = 0 \quad \text{in } \mathbb{R}^2 \\
& \quad \text{and} \quad u = f \text{ at } \partial \mathbb{D} \\
\end{align*} \]

\[ \begin{align*}
\text{(a)} & \quad \text{max} \quad -\frac{1}{2} \int_{\mathbb{R}^2} |v|^2 \, dx \\
& \quad \text{subject to} \quad \Delta v = 0 \quad \text{in } \mathbb{R}^2 \\
& \quad \text{and} \quad v = f \text{ at } \partial \mathbb{D} \\
\end{align*} \]

These problems have the same relationship to each other as our \( P \) and \( D \) in linear programming.
(a) If $x$ is admissible for $D$ and $u$ is admissible for $P$, then

value of $D$ at $x$ \leq value of $P$ at $x$.

Equality holds when $u^*$ solves $P$ and $x^*$ solves $D$.

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

To see (a), expand

$$\int_{S^2} \frac{1}{2} |x|^2 - \frac{1}{2} |u|^2 \, dx \geq 0$$

To see

$$\int_{S^2} \frac{1}{2} |x|^2 + \frac{1}{2} |u|^2 - \langle x, 7u \rangle \, dx \geq 0,$$

then use $\det \sigma = 0$ on $S^2$, \( \sigma_{ij} = \delta_{ij} \) to get

$$-\frac{1}{2} \int_{S^2} |x|^2 \leq \frac{1}{2} |u|^2 - \int_{S^2} \frac{du^*}{dx}$$

To see (b), observe that if $u^*$ solves
and \( \sigma^* = 72 \mu^+ \) then (from p. 51 (c))
\( \sigma^* \) solves (65). [Since all the maps are equalities in the disc of (67).] Note 8
thus \( \sigma^* \) is the only soln of (65), since the
1st \( \int \mu^2 \) is strictly convex.

Note 9: in this case the choice of the optimal choices

\[
\begin{align*}
div \sigma^+ &= 0 \quad \text{in } \mathbb{R}^2 \\
\sigma^* &= 72 \mu^+ \\
\sigma^*, \mu^+ &= f \quad \text{at } \partial \Omega
\end{align*}
\]

amounts to a requirement of the EL eqn for
\( \Omega \) (\( du^+ = 0 \) in \( \mathbb{R}^2 \), \( du^+ = f \) at \( \partial \Omega \)). [When
the EL of \( \Omega \) makes sense, the complementary
weakness under is always equivalent to
the EL eqn.]

Deja vu? : I skipped over the question:
what is the proper functional space for \( \sigma^* \)?
[Even then, if \( \sigma \) is smooth — but
hard to know there is an optimal \( \sigma \). That's
smooth?]. Given form of \( \sigma^* \), natural space
to look (e.g. using the Direct Method) is

\[
X = \{ \sigma : \int_{\Omega} \mu^2 < \infty \quad \text{and} \quad div \sigma = 0 \ \text{in } \mathbb{R}^2 \}
\]
Does the constraint "\( \sigma \cdot n = f \) at \( \partial \Omega \)" make sense? Answer is yes: there's a cent's map
\[ X \rightarrow H^{-1/2}(\partial \Omega) \]
taking \( \sigma \rightarrow \sigma \cdot n \) when \( \sigma \) is smooth.

and Green's formula holds in the sense that
\[ \int_{\partial \Omega} \sigma \cdot n \cdot u = \int_{\Omega} \left< \nabla \sigma, \nabla u \right> + \int_{\Omega} \sigma \cdot \nabla u \] for all \( u \in H^1(\Omega) \).

Hence, \( H^{-1/2}(\partial \Omega) \) is the dual of \( H^{1/2}(\partial \Omega) \) in \( L^2 \)-inner product.

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

The special case \( \partial \Omega = \mathbb{R}^2 \) is more elementary.

Since \( \text{div} \, \sigma = 0 \) \( \Rightarrow \sigma = (\nabla \phi)^\perp \), and \( \sigma \cdot n \perp \sigma \cdot n \) in \( \partial \Omega \).

Digression 2: When doing numerical calculations by finite element methods, it can be difficult to know how good an approximation of the value you have obtained. A "primal-dual" method studies \( \sigma \) and \( n \) simultaneously.
Advantage of this is evident from:
Lemma: If $\tilde{u}$ is admissible for $\Sigma$ and $\tilde{v}$ is admissible for $\Pi$ and

\[(\text{value of } \Pi \text{ at } \tilde{u}) - (\text{value of } \Gamma \text{ at } \tilde{v}) \leq 5\]

Then

\[\frac{1}{2} \left( \left| 17\tilde{u} - 7\tilde{u}^2 \right|^2 \right) \leq 5 \quad \text{and} \quad \frac{1}{2} \left( \left| 10\tilde{v} - 5\tilde{v}^2 \right|^2 \right) \leq 5\]

where $\tilde{u}$, $\tilde{v}$ are the values of $\Sigma$ and $\Pi$.

(Proof: exercise. Note the similarity to Exercise 6 at the end of Lecture 1.)

How to find dual problems systematically?

Central idea: a convex optimization can be expressed as a min max. Switching the min and the max gives the dual problem. (There can be more than one min/max representation of a given convex problem; different choices may lead to slightly different "dual" plans.)

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

\[(P) \quad \min_{\tilde{u}} \int W(\tilde{z}(\tilde{u})) - \int \frac{\tilde{u}}{2} \]
with \( W(\xi) \) convex and \( \int f d\xi = 0 \).

Key pt: \( W \) convex \( \Rightarrow \) its graph is the envelope of its supporting hyperplanes.

\[ \iff \quad W(\xi) = \sup_{\gamma} \langle \gamma, \xi \rangle - W^*(\gamma) \]

where \( W^*(\gamma) \) (the "Fejerian transform" of \( W \)) is defined by

\[ W^*(\gamma) = \sup_{\xi} \langle \xi, \gamma \rangle - W(\xi) \]

So:

\[ \min_u \int_{\mathbb{R}} W(7u) \, dx = \min_u \sup_{\gamma} \int_{\mathbb{R}} \langle \xi, \gamma \rangle - W(7\xi) \, dx \]

\[ = \min_u \max_{\gamma} \left\{ \int_{\mathbb{R}} \langle \xi, \gamma \rangle - W^*(\gamma) \, dx \right\} \]

Claim: we can switch \( \min \) and \( \max \). (Return to this...)

\[ \min \sup \]

\[ \max \inf \]

\[ \min \inf \]

\[ \max \sup \]
soon). Then above becomes
\[ \max_x \min_u \left[ \int (u_0 - f) u \, du \right] = \int \left( d\omega(\sigma \cdot u + W_x) \right) \, dz. \]

(we transformed by parts in the 1st line above & observed that the min w.r.t. \( u \) is -∞ unless \( d\omega(\sigma) = 0 \) \& \( \sigma \cdot u = f \) to get 2nd line).

Why should \( \min \max = \max \min \)? For simplicity, it is trivial, with no structural hypothesis:

**1st attempt:** \[
\min_y F(x,y) \leq F(x,y_0).
\]
\[
\Rightarrow \max_x \min_y F(x,y) \leq \max_x F(x,y_0) \]
\[
\Rightarrow \max_x \min_y F(x,y) \leq \min_x \max_y F(x,y).
\]

**2nd attempt:** if \( d\omega(\sigma) = 0 \) and \( \sigma \cdot u = f \) then internal to the plume inequality
\[ W(\nabla u) \geq <\nabla u, \sigma> - W^*(\sigma) \]
\[ \int W(\nabla u) - \int u \cdot f \geq -\int W^*(\sigma) \]

Either way, we see that for every \( \sigma \) (admissible for \( \Phi \)), we get a lower bd for \( \Phi \).

The fact that we get equality, i.e.,
\[ \max f = \min \Phi, \]

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 convexity on \( F \) (typically \( \sigma \)-concave in \( y \),

convex in \( x \), and a little more - see eg.

book by Ekeland + Temam).

But: if \( \Phi \) 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 \( \infty \) ( \( W(\xi) \sim |\xi|^p \) as \( |\xi| \to \infty \) and \( dW/|\xi| \sim |\xi|^{p-1} \) as \( |\xi| \to \infty \)). Then

minimizer is a weak soln of EL eqn

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

Solv of dual prob is then \( \sigma^* = \frac{\partial W}{\partial \nabla u} \big|_{u=\sigma^*} \).
Since each of this set $D$ goes into $B$, we get an explicit $P$ that \( \inf B = \sup D \) without need for any minimax theorem. (Why is value of $\inf B$ at this $u$ equal to $\inf B$? Exercise.)

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

Discussion: In discussing the "Direct Method" we used a theorem from real analysis to show that
\[
E[u] = \frac{1}{2} \int W(u) \, dx
\]
is lower semicontinuous under weak convex of $u$ (if $W$ is convex with $p$th power growth $1 < p < \infty$). The Fenchel transform gives a different, more intuitive proof of lower semicontinuity. In fact,
\[
E[u] = \sup \int \frac{\partial (u, v) - W^*(v)}{2} \, dx
\]
and if $v$ is fixed then
\[
u \rightarrow \frac{\partial (u, v) - W^*(v)}{2} \, dx
\]
is continuous under weak convergence (in $W^{1,p}$) since
it is linear in \( u \). So

\[ E[u] = \text{max of fields that are cnt's (under w.e convergence)} \]

is lower semicontinuous under w.e. convergence.

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

Let \( \lambda_0 = \text{1st Dirichlet eigenvalue of a domain } S \subset \mathbb{R}^n \)

\[
= \min_{u=0 \text{ at } S^c} \frac{\int_S 17x_1^2 \, dx}{\int_S u^2 \, dx} \frac{\int_S u^2 \, dx}{\int_S \, dx}.
\]

\[
= \min_{u=0 \text{ at } S^c} \frac{\int_S 17x_1^2 \, dx}{\int_S \, dx}.
\]

Upper lds are easy (consider any \( u \)). How about a scheme for proving lower bounds?
Step 1: Suffice to consider \( u \geq 0 \), since replacing \( u \) by \( |u| \) leaves both \( \frac{1}{2} u^2 \) and \( \frac{1}{2} u^2 \) unchanged. (Exercise.)

Step 2: Let \( p = u^2 \) (i.e. let \( u = \sqrt{p} \)) and write the delta of \( \Delta u \) in terms of \( p \):

\[
\Delta u = \min_{p \geq 0} \int_0^1 \frac{1}{2} p \sqrt{p} \, dx
\]

\[
\frac{p}{2} = 1 \quad \text{if} \quad p \geq 0 \quad \text{on} \quad [0,1].
\]

\[
p = 0 \quad \text{at} \quad 2\pi.
\]

Step 3: Observe that \( \frac{\pi^2}{4} + \) is a convex for \( (\xi, t) \) in fact:

\[
\frac{\pi^2}{4} + t = \max_{\theta} \langle \sigma, \xi \rangle - H(\sigma)^2
\]

So

\[
\Delta = \min_{p \geq 0} \max_{\theta} \int_0^1 \langle \sigma, \xi \rangle - H(\sigma)^2
\]

\[
\frac{p}{2} = 1 \quad \text{if} \quad p \geq 0 \quad \text{on} \quad [0,1].
\]

\[
p = 0 \quad \text{at} \quad 2\pi.
\]
**Step 4** Switch min to max to get a dual plan:

\[ \lambda_0 = \max_{\phi} \min_{\psi} \int_{\Omega} \left[ \psi \Delta \phi + \phi \psi \right] \, dx \]

subject to:

- \[ \int_{\Omega} \phi = 1 \]
- \[ \phi \geq 0 \text{ in } \Omega \]
- \[ \phi = 0 \text{ at } \partial \Omega \]

\[ = \max_\mu \mu \left[ \text{div}^2 \phi + |\phi|^2 \right] \geq \mu \text{ (constant!)} \]

\[ = \text{largest constant } \mu \text{ st } \exists \text{ vector field } \phi \text{ on } \Omega \text{ with } \text{div}^2 \phi + |\phi|^2 \leq -\mu \text{ ptwise.} \]

**Step 5:** Is the max/min right? Sure! Observe that \[ \max_{\psi} \left< \Delta \phi, \psi \right> = |\phi|^2 \] is achieved when \[ \phi \equiv \text{ const.} \]

So, best \( \phi \) is \[ \frac{1}{\sqrt{p}} \phi \] where \( p = u^2 \) and \( u \) is 1st Dirichlet exponent. Direct calculation \[ \text{This } \phi \text{ is admissible for proposed dual plan and achieves its optimal value! (Exercise!)} \]

**Suggested exercises**

1. [This plan could have been at the end of]
Lecture 1.7 In Lecture 1 we used the Direct Method of the Calc of Var to show that if $W(\varepsilon)$ is convex with

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

then there exists $u \in W^p(\Omega)$ such that

$$\min_{u \in \partial^{\Omega}} \int \frac{1}{2} W(\nabla u) + f u \, dx$$

Show that if $W$ is differentiable and

$$\left| \frac{\partial W}{\partial \varepsilon} \right| \leq C |\varepsilon|^p$$

as $|\varepsilon| \to \infty$

Then the minimizer satisfies the EL eqn in the sense that

$$\int \frac{1}{2} \nabla (\nabla u), \nabla v + f v \, dx = 0$$

for any $v \in W^p_0(\Omega)$ at $v = 0$ at $\partial \Omega$. The main task is to prove by putting the $\frac{\partial}{\partial \varepsilon}$ under the integral in

$$\frac{\partial}{\partial \varepsilon} \int \frac{1}{2} W(\nabla u + \varepsilon \nabla v) + f(u + \varepsilon v) \, dx$$

(2) On pg 2.2 I mentioned, as non-smooth example,

$$\min \int W(\nabla u) \, dx \text{ with}$$

$$u = g \text{ on } \partial \Omega$$
\[ W(\mathbf{u}) = \begin{cases} 
21\mathbf{u} & |\mathbf{u}| \leq 1 \\
1 + 17\mathbf{u}^2 & |\mathbf{u}| > 1 
\end{cases} \]

What is the convex dual to this variational problem?

3. Show that the dual of

\[
\min_{\mathbf{u} \in \mathbb{S}} \int_{\mathcal{D}} \frac{1}{2} 17\mathbf{u}^2 \\
\text{subject to } \mathbf{u} = \mathbf{g} \text{ at } \partial \mathcal{D}
\]

is

\[
\max_{\phi \in \mathcal{V}_0} \int_{\mathcal{D}} (\phi, \mathbf{n}) \phi - \frac{1}{2} \int_{\mathcal{D}} |\nabla \phi|^2
\]

4. Do the exercise suggested on pg 1.9

5. Do the exercise suggested near the top of pg 2.13