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

1) Suppose we're interested in the min value of a convex optimization, e.g.
   \[ \min \int W(x) \, dx \]
   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.

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

   \[ \min \int f(x) \, dx \]
   \[ g(x) = 1 \]
   \[ h(x) = 0 \]
   \[ \text{at } x = 25 \]

   The convex dual provides rec+mult
conds for optimality (playing a role analogous to the Euler-Lagrange eqn of a smoother convex prob).

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

\[
\begin{align*}
(\mathcal{P}) \quad \min & \quad \sum_{j=1}^{n} a_{j}x_{j} \\
\text{subject to} & \quad \sum_{j=1}^{n} a_{j}x_{j} \geq b_{i} \quad i \leq m \quad \quad x \geq 0
\end{align*}
\]

We can derive "trivial lower bd" on opt l value
by taking the combo of constraints: if \( y_{i} \geq 0 \)
and \( \sum_{i=1}^{m} a_{i}y_{i} \leq c_{j} \) then

\[
y_{i} \frac{\sum_{j=1}^{n} a_{j}x_{j}}{y_{i}} \geq b_{i} \quad \Rightarrow \quad \sum_{j=1}^{n} c_{j}x_{j} \geq \sum_{i=1}^{m} b_{i}y_{i}
\]

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

\[
\begin{align*}
(\mathcal{D}) \quad \max & \quad \sum_{i=1}^{m} b_{i}y_{i} \\
\text{subject to} & \quad \sum_{i=1}^{m} a_{i}y_{i} \leq c_{j} \quad y_{i} \geq 0
\end{align*}
\]

Duality theorem of lin prog 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
thus is not trivial: see P. Lax's linear book for
an attractive approach close to spirit of this lecture. But any new proof text will have a proof; I like the book by Christol.

Note: if \( \mathbf{x}^* \) solves dual + \( \mathbf{x} \) solves primal
then (by duality theorem) \( \Sigma_i c_i x_i^* = \Sigma_i b_i y_i^* \).

Examining the column on page 12 we see that certain relations must hold:

\[
\forall i: \quad y_i^* > 0 \quad \sum_j a_{ij} y_j^* \geq b_i \quad \text{with equality in at least one of the two}
\]

\[
\forall j: \quad x_j^* > 0 \quad \sum_i a_{ij} x_i^* \leq c_j \quad \text{with equality in at least one of the two}
\]

Thus: existence of solution \( \mathbf{y}^* \) to these "complementary slackness conditions" replaces the Euler-Lagrange eqns (as for a smooth convex problem, any cost \( \mathbf{p}^* \) is a minimum).

Here is a basic but typical example of two problems in duality. Suppose \( f: \mathbb{R}^n \to \mathbb{R} \) satisfies \( \int f 0 = 0 \). Consider

\[
\min_{\mathbf{u}} \int_0^1 \frac{1}{2} |\mathbf{u}'|^2 - \int f \mathbf{u} \mathbf{d}s
\]
\[ \text{(10) } \quad \text{max} \quad - \frac{1}{2} \int_{\Omega} |\nabla u|^2 \, \text{d}x = 0 \quad \text{on } \Gamma \]
\[ \frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma \]

[Remark: we use here the fact that \( \text{div} \sigma = 0 \), \( \sigma \in L^2 \).]

On has a well-defined trace on \( \partial \Omega \), for which Green's theorem holds:
\[ \int_{\partial \Omega} (\sigma \cdot n) u = \int_{\Omega} \nabla u \cdot \nabla v \, \text{d}x + \int_{\partial \Omega} \sigma \cdot n \, u \, \text{d}s \]

for all \( u \in H^1(\Omega) \). In general, \( \sigma \cdot n \in H^{-1/2}(\partial \Omega) \) is dual to \( H^{1/2}(\partial \Omega) \), since \( H^{1/2}(\partial \Omega) \) is exact space of traces of \( H^1(\Omega) \).

Special case \( \Omega = \mathbb{R}^2 \) is easiest since \( \text{div} \sigma = 0 \Rightarrow \sigma \in L^2 \) and \( \sigma \cdot n = \text{grad} \varphi \) on \( \partial \Omega \).

The primal and dual have same relation as before:

1. if \( \{ u \text{ admissible} \} \) and \( \{ \sigma \text{ admissible} \} \) then
\[ \text{value of } \Phi \text{ at } \sigma \leq \text{value of } \Psi \text{ at } u \]

2. and equality holds when \( u \) solves \( \Phi \) or solves \( \Psi \).

Proof in this case is elementary: to see 1,
\[ \int_{\Omega} \frac{1}{2} |\nabla u|^2 \, \text{d}x \geq 0 \]
to see \( \frac{1}{2} \| \sigma \|^2 + \frac{1}{2} \| 7u \|^2 \leq \langle \sigma, 7u \rangle \geq 0 \).

Now use \( d \sigma = 0 \), \( \sigma \cdot n = f \) to get

\[
- \frac{1}{2} \int \| \sigma \|^2 \geq \frac{1}{2} \int 17u^2 - \frac{1}{2} \int f 252
\]

To see (2), observe that if \( u^* \) solves primal

\[ + \hat{\sigma} = 7u^* \] then preceding inequalities are equal.

Thus \( \hat{\sigma} \) in this case the exact equality

\[
d \sigma = 0 \quad and \quad \hat{\sigma} = 7u
\]

\( \sigma \cdot n = 0 \)

are just a rewrite of the EL eqns in \( \hat{\sigma} \).

Rule: When doing numerical calculations by

finite element method it can be difficult to

know how good an approach you have obtained.

A "primal-dual" method can study \( P + \hat{P} \)

simultaneously. Lemma 3 if \( \hat{\sigma} \) is admissible

in \( \hat{P} + \hat{u} \) is admissible to \( P \) and

\[ (\text{value of } P \text{ at } \hat{u}) - (\text{value of } \hat{P} \text{ at } \hat{\sigma}) < \delta \]

Then \( \frac{1}{2} \int \| \hat{u} - 7u^* \|^2 \leq \delta \) and \( \frac{1}{2} \int \| \hat{\sigma} - \hat{\sigma} \|^2 \leq \delta \)
where $u^*$ and $v^* = T u^*$ are the values of $P + D$.

(Part: exercise.)

How could we have found the dual problem systematically? Any dual pairs are assoc to "saddle pt" wold plans (ie to switch min & max) Explain by considering the more general problem

$(P) \quad \min \; \frac{1}{2} \sum_{ij} W(x_{ij}) - \sum_i u_i f_i$

with $W(\xi)$ convex.

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

$\iff W(\xi) = \sup_\gamma \langle \gamma, \xi \rangle - W^*(\gamma)$

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

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

slope $\gamma$

$y$-intercept determines $W^*(\gamma)$. 
So:

$$\min_u \int W(\sigma u) - \int u \cdot f \quad \min_u \max_{\sigma(x)} \int <\sigma, \nabla u> - W(\sigma) - \int u \cdot f$$

$$= \min_u \max_{\sigma(x)} \int \sigma \cdot (\nabla u - f \cdot u + \int (-\nabla \sigma) u - W(\sigma))$$

Claim: we can switch \( \min + \max \) (Refer to this soon)

$$= \max_{\sigma(x)} \min_u \int (\sigma \cdot \nabla u - f \cdot u - \int (\nabla \sigma) u + W(\sigma))$$

$$= \max_{\sigma(x)} \int W^*(\sigma) \quad \min_{\sigma(x)} \\sigma \cdot \nabla u = f$$

since if \( \nabla \sigma \neq 0 \) or \( \sigma \cdot \nabla u \neq f \) then \( \min \) over \( u \)

would be \(-\infty\).

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

1st view: \( \min_y F(x,y) : F(x,y) \leq F(x,y_0) \)

\( \Rightarrow \max_x \min_y F(x,y) \leq \max_x F(x,y_0) \)

\( \Rightarrow \max_x \min_y F(x,y) \leq \min_x \max_y F(x,y) \)
2nd script: suppose $\nabla \sigma = 0$, $\sigma \cdot n = f$.

Interpret

$$W(v) \geq <v, \sigma> - W^*(\sigma)$$

to get

$$\int \frac{1}{2} W(v) - \int u \cdot f \geq -\int W^*(\sigma)$$

The fact that we get equality (not inequality) is nontrivial in many cases. But if $Q = \sigma \cdot n$ has a clearly defined EL $v$ such that it will give us a direct proof. In present setting: if $W$ is smooth enough that

$$\partial_t (\frac{\partial W}{\partial \sigma}) = 0$$

has a solution, then we can take $\sigma = \frac{\partial W}{\partial \sigma}$.

(Note: when $W(v) = \frac{1}{2} |v|^2$, $W^*(\sigma) = \frac{1}{2} |\sigma|^2$.)

Plan for rest of lecture:

a) give a more subtle example that still involves linear PDE
b) discuss some examples involving $L^1 - L^\infty$ duality. (where a conventional Euler-Lagrange eqn can't be written)

Here's example (as): Let $\lambda_0$ be $1^{st}$ Dirichlet eigenvalue of domain $\Omega$:

$$\lambda_0 = \min_{u \geq 0 \text{ at } \partial \Omega} \frac{\int \frac{u'^2}{2}}{\int |u|^2} = \min_{u \geq 0 \text{ at } \partial \Omega} \frac{\int \frac{17u'^2}{2}}{\int |u|^2} = \frac{\int \frac{17u'^2}{2}}{\int |u|^2} = 1$$

Upper bd is easy (consider any $u$). How about a scheme for proving lower bounds? 

**Step 1**. Suff+ to consider $u \geq 0$, since replacing $u$ by $|u|$ leaves both $\int \frac{17u'^2}{2}$ and $\int u^2$ invariant. (Exercise)

**Step 2**. Let $p = u^2$ (i.e. let $u = \sqrt{p}$) + write definition of $\lambda_0$ in terms of $p$:

$$\lambda_0 = \min_{p \geq 0 \text{ in } \Omega} \frac{\int \frac{17p^2}{2p}}{\int p} \cdot dx$$

$$\int p \cdot dx = 1$$

$p \geq 0 \text{ in } \Omega$

$p = 0 \text{ at } \partial \Omega$
Step 3: Observation: the function \( f(x) = \frac{1}{2} x^2 / 4 + 1 \) is convex for \( x \in (0, 1) \). In fact,

\[
\frac{1}{4} x^2 = \max_\sigma \langle \sigma, \frac{\sigma}{\|\sigma\|} \rangle - \frac{1}{2} ||\sigma||^2
\]

So, \( \lambda_0 = \min_\sigma \max_\rho \int_0^1 \frac{1}{2} \langle \sigma, \frac{\sigma}{\|\sigma\|} \rangle - \rho ||\sigma||^2 \)\)

\( \sigma \geq 0 \)

\( \rho \geq 0 \)

\( \rho = 0 \) at \( 2\pi \)

Step 4: Proceed as usual: switch min/max, and use min over \( \rho \) to get constants in \( \sigma \)

\( \lambda_0 = \max_\sigma \min_\rho \int_0^1 \rho \cdot \sigma - \int_0^1 \rho (\rho \sigma + 10^2) \)

\( \sigma \geq 0 \)

\( \rho \geq 0 \)

\( \rho = 0 \) at \( 2\pi \)

\[ \lambda_0 = \max_\sigma \mu \]

\( \rho \sigma + 10^2 = -\mu \text{ constant} \]

\[ \lambda_0 = \text{largest constant } \mu \text{ of vector field } \sigma \]

\( \sigma \cdot \rho \sigma + 10^2 = -\mu \)

Step 5: Is the max/min right? Sure!

\( \max_\sigma \langle \sigma, \frac{\sigma}{\|\sigma\|} \rangle - \frac{1}{2} ||\sigma||^2 \) is achieved when \( \sigma = 2 \pi \).
So best $\sigma$ is $\frac{1}{2p} \sqrt{p}$ where $p = u^2 + v^2$ in 1st Dirichlet example. Direct calculus is admissible for dual problem and achieves its optimal value. (Exercise.)

Here's example (6): What can we say about

\[
\min \left\{ \| \sigma \|_{1,\infty} s.t. \ \text{div} \ \sigma = 1 \text{ on } \Omega \right\}
\]

where (to fix ideas) $\Omega$ is a bounded domain in $\mathbb{R}^2$.

(Interpretation: it's raining uniformly on $\Omega$.
How can rain flow to below with least possible local accumulation?)

1st observation: it's equivalent to solve

\[
\min \lambda \quad \text{subject to} \quad |\sigma| \leq 1, \ \text{div} \ \sigma = \lambda \text{ (constant)}
\]

Since if $\lambda_{\max}$ is optimal value of (**) then

\[
\text{div} \ \sigma = \lambda \text{ (const.) } \Rightarrow \lambda \leq \lambda_{\max}
\]

$|\sigma| \leq 1$
So \( \text{div} \left( \frac{\mathbf{F}}{x} \right) = 1 \Rightarrow \frac{1}{x} \geq \frac{1}{\lambda_{\max}} \)

\[ \frac{1}{x} = \frac{1}{\lambda_{\max}} \]

so \( \frac{1}{\lambda_{\max}} \) is opt'1 value for \((\ast)\).

Identify dual of \((\ast\ast)\) by usual argument:

\[
\max_{10 \leq 1} \min_{u=0 \at \Omega} \int_{\Omega} \frac{\mathbf{a}^T \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{x}} = \min_{u=0 \at \Omega} \max_{10 \leq 1} \int_{\Omega} \mathbf{a}^T \cdot \mathbf{u}
\]

\[
\int_{\Omega} u = 1
\]

\[
\text{min over } u = -\infty
\]

\[\text{unless } \text{div } \mathbf{a} = \lambda \text{ constant} \]

\[
\downarrow
\]

\[
\max_{10 \leq 1} \lambda
\]

\[
\text{div } \mathbf{a} = \lambda \text{ constant}
\]

\[
\downarrow
\]

\[
\text{min}_{u=0 \at \Omega} \int_{\Omega} \mathbf{a}^T \cdot \mathbf{u}
\]

\[
\int_{\Omega} u = 1
\]

Is \( \mu u = \mu u \), inequality is elementary as usual:

\[
10 \leq 1 \Rightarrow \int_{\Omega} \mathbf{a}^T \cdot \mathbf{u} \leq \int_{\Omega} \mu u
\]

\[
u = 0 \at \Omega
\]

\[
\text{div } \mathbf{a} = \lambda
\]

\[
\int_{\Omega} u = 1
\]
However, equality is not simple to prove. Moreover, optimal u is rather singular — it's the characteristic function of a set (see below). That max min = min max can be proved using results from books by Ekeland + Temam (Convex Analysis + Variational Problems).

Similar L^1-L^\infty duality problems arise in plasticity & other areas of mechanics; see e.g. book by Duvaut + Lions (Inequalities in Mechanics + Physics).

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

\[
(\text{**}) \quad \min_{\text{u=0 on } \partial D} \int_D |\nabla u| = \min_{D \subseteq \mathbb{R}^2} \frac{\text{length}(\partial D)}{\text{Area}(D)}
\]

\[
\int_D u = 1
\]

... and of a geometry problem!

Example for D = square:

![Square and circle graphic]

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

Sketch proof of (**). A key ingredient is the
\[ \int f(x) \, dx = \int \left( \int f(x) \, dx \right) \, dt \]

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

Step 1: LHS of (***): \[ \min_{u=0}^{u=2\pi} \frac{1}{u} \int u \, dx \]

(easy)

Step 2: May assume \( u \geq 0 \) (since replacing \( u \to 1/u \) leaves \( \int u \, dx \) invariant and increases \( \frac{1}{u} \)).

Step 3: For \( u \geq 0 \), \[ \int u \, dx = \int_0^\infty \text{Area } \{ u \geq t^2 \} \, dt \] since \[ \int u \, dx = \int_0^\infty \int_0^{u(x)} 1 \, dt \, dx \]

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

\[ \int \int u \, dx = \int_0^\infty \text{Length } \{ u = t^2 \} \, dt \]

(coarea formula with \( f=1 \)). So if RHS of (***), has min \( \times \) them
length $\|u\| = \frac{1}{\lambda}$ \geq \alpha \text{ area } \Omega \|u\| = \frac{1}{\lambda}$ for all $t$

\[\Rightarrow \int_{\Omega} |\nabla u| \geq \alpha \|u\| \int_{\Omega} \]

\[\Rightarrow \text{ LHS of (4.4) } \geq \alpha.\]

For more discussion of closely related plans see recent paper by G. Strong, "Maximum flow and minimum cuts in the plane." (J. Global Optim., in press. There's a preprint on Strong's website.) I'll take just one item from there: a very efficient cut (due to Greenspan) y

Cheeger's inequality: if $\lambda = \min_{D \subset \Omega} \text{ eigenvalue}$

\[h = \min_{D \subset \Omega} \frac{\text{Length}(\partial D)}{\text{Area}(D)}\]

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

Proof: From prior discussion, $f = \mathbb{1}_{\Omega \setminus \text{dist} \Omega, \partial D = \Omega}$

Let $u_0$ be 1st dir. eigenfun. Then
\[ \Phi \leq 2 \left( \frac{\int u_0^2}{\int u_0^2} \right)^{1/2} \]

\[ = 2 \left( \frac{\int u_0^2 (\int u_0^2)^{1/2}}{(\int u_0^2)^{1/2}} \right)^{1/2} \]

\[ = 2 \lambda_0^{1/2} \]

**Some exercises:**

1) Show that the convex dual of

\[ \min_{u = u_0 \text{ at } \partial \Omega} \int \frac{1}{2} \| u \|^2 \]

is

\[ \max_{\sigma \cdot n = F \text{ at } \partial \Omega} \int \frac{1}{2} \| \sigma \|^2 \]

2) Show that the variational problem

\[ (\mathcal{P}) \min_{\sigma \cdot n = f \text{ at } \partial \Omega} \int_{\Omega} |\sigma| \]

is
(25) \[ \max_{u} \int_{\Omega} u f \, dx - \int_{\partial \Omega} u F \, ds \]

with \[ 17u_{\infty} \leq 1 \]

are a dual pair, i.e. \( \int_{\Omega} F f \, dx = \int_{\partial \Omega} F \cdot n \). How should \( \sigma \) and \( \varepsilon u \) be related if equality is to hold?

\[ \text{find } \sup_{\sigma \in \mathbb{R}^n} \left\{ \frac{\varepsilon}{2} \| \sigma - 10 \| \right\} = \begin{cases} 0 & \text{if } |\varepsilon| \leq 1, \\ +\infty & \text{if } |\varepsilon| > 1. \end{cases} \]

Rule: if \( \Omega \subseteq \mathbb{R}^2 \) and \( F = 0 \) then \( Q \) can be solved explicitly in simple cases using the co-area formula. Why?

3) When we study homogenization we'll consider a periodic conductivity \( a(x) \), and we'll learn that the above effective conductivity

\[ <a_{\text{eff}}, \phi> = \min_{\phi \text{ periodic}} \int_{\Omega} <a(x) (\phi + \phi_0), \phi + \phi_0> \]

\( \phi_0 \) is in general a matrix; \( \phi \) denotes the average of a periodic function. Show using duality that
\[ \langle \text{eff } \xi, \xi \rangle = \max_{\delta \sigma = 0} \int_{\sigma \text{ periodic}} 2\delta, \xi \rangle - \langle \text{eff } \xi, \xi \rangle = \min_{\delta \sigma = 0} \int_{\sigma \text{ periodic}} \langle \xi^{-1}(x)\sigma, \sigma \rangle \]

This can alternatively be written as

For any \( \xi, \sigma \in \mathbb{R}^n \). Optimizing over \( \xi \), we get

\[ \langle \xi^{-1} \eta, \eta \rangle = \min_{\delta \sigma = 0} \int_{\sigma \text{ periodic}} \langle \xi^{-1}(x)\sigma, \sigma \rangle \]