Lecture 5: Convex Duality

Goals:

1. Brief discussion of linear programming
2. Convex duality in P.D.E. setting
3. Derivation of dual problem
4. Calibration method.

1. Motivation

a.) In the homework, we found an upper bound for the problem

$$\inf_{f\in A} I[f] = \inf_{f\in A}\left(\int_0^1 (f'(x)^2-1)^2\,dx + \varepsilon^2\int_0^1 f''(x)^2\,dx\right) \leq \min\{1, C\varepsilon\}.$$

The upper bound is easy to ~~solve~~ find — we just make a good guess. Finding a lower bound that scales in the same manner is more tricky. For problems of the form

$$\cancel{\inf_{f\in A} I[\nabla f]}\quad \inf_{f\in A} I[f] = \inf_{f\in A}\int_\Omega L(\nabla f)\,dx,$$

the convex dual provides a systematic approach.

b.) Sometimes we have convex but non-smooth Lagrangians:

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

The convex-dual provides necessary + sufficient conditions for optimality.

2. Linear Programming

Consider the primal problem

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

Lets derive a lower bound. If

$$y_i \geq 0 \text{ and } \sum_{i=1}^m a_{ij}y_i \leq c_j$$

(Notice index being summed over)

Then,

$$y_i \sum_j^n a_{ij} x_j \geq b_i \cdot y_i$$

$$\Rightarrow \sum_{j=1}^n \sum_{i=1}^m y_i a_{ij} x_j \geq \sum_{i=1}^m b_i y_i$$

$$\Rightarrow \sum_{j=1}^n c_j x_j \geq \sum_{i=1}^m b_i y_i .$$

There fore,

$$\boxed{\min_{\substack{\sum_{j=1} a_{ij} x_j \geq b_i \\ x_j \geq 0}} \sum_{j=1}^n c_j x_j \geq \max_{\substack{\sum_{i=1} a_{ij} y_i \leq c_j \\ y_i \geq 0}} \sum_{i=1}^m b_i y_i}$$

The maximum problem is called the dual problem

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

The duality theorem of linear programming says
max (D) = min (P).

Important - If y^* solves D and x^* solves P then

$$\sum_j c_j x_j^* = \sum_i b_i y_i^* .$$

$$\Rightarrow \sum_{j=1}^n \sum_{i=1}^m y_i^* a_{ij} x_j^* = \sum_{i=1}^m b_i y_i^* = \sum_{j=1}^n c_j x_j^*$$

Optimizing gives

$$\boxed{\begin{array}{l} \forall_i, \\ \quad y_i^* \geq 0 \text{ and } \sum_j a_{ij} x_j^* \geq b_i \text{ (with equality in at least one)} \\ \forall_j, \\ \quad x_j^* \geq 0 \text{ and } \sum_i a_{ij} y_i^* \leq c_i \text{ (with equality in at least one).} \end{array}}$$

3. Example

Consider:

$$(P) \quad \min \int_{\Omega} \frac{1}{2}|\nabla f|^2 dx - \int_{\partial\Omega} f g \, dx$$

$$(D) \quad \max_{\substack{\nabla\cdot\sigma = 0 \text{ in } \Omega \\ \sigma\cdot\vec{n} = g \text{ on } \partial\Omega}} -\frac{1}{2}\int_{\Omega} |\sigma|^2 dx$$

Note: a.) Value of $D <$ Value of P

b.) Equality holds if f minimizes (P) and σ maximizes (D).

~~proof~~:

a.) $\int_{\Omega} \frac{1}{2}|\sigma - \nabla f|^2 \geq 0$

$$\Rightarrow \frac{1}{2}\|\sigma\|_{L^2}^2 - \langle \sigma, \nabla f \rangle + \frac{1}{2}\|\nabla f\|_{L^2}^2 \geq 0$$

$$\Rightarrow \frac{1}{2}\|\sigma\|_{L^2}^2 - \int_{\partial\Omega} f \cdot \sigma\cdot\vec{n} + \int_{\Omega} \cancel{(\nabla\cdot\sigma) f}^{\,0} dx + \frac{1}{2}\|\nabla f\|_{L^2}^2 \geq 0$$

$$\Rightarrow (P) \geq (D).$$

b.) To prove (b), suppose f^* solves (P) then setting $\sigma^* = f^*$ we get equality of the bounds. (This is the only solution by strict convexity of D). ~~Euler Lagrange equations are then~~ Equation satisfied by σ is then

$$\boxed{\begin{array}{ll} \nabla\cdot\sigma = 0 \text{ in } \Omega, & \sigma = \nabla f \\ \sigma\cdot\vec{n} = g \text{ on } \partial\Omega & \end{array}}$$

4. Legendre - Fenchel Transform

Focus on

$$(P) \quad \min_f \int_{\Omega} L(\nabla f) - \int_{\partial\Omega} f\cdot g \, dx.$$

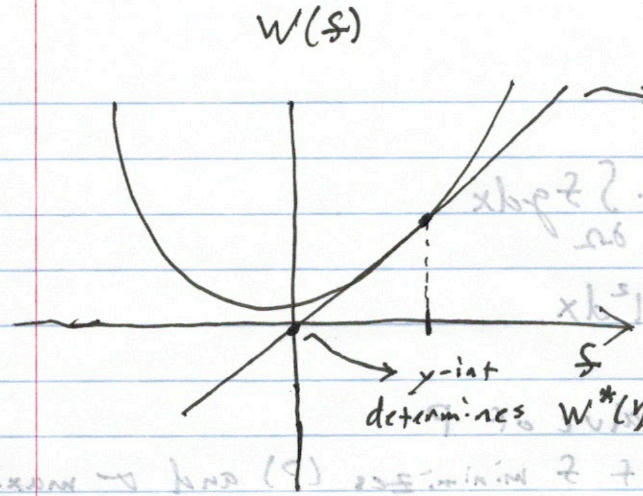
with L convex and $\int_{\partial\Omega} g \, ds = 0$.

Key point: L convex $\Longleftrightarrow$ sup of supporting hyperplanes

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

where W^* is the Fenchel transform defined by

$$W^*(\eta) = \sup_f \{ \langle \eta, f \rangle - W(f) \}$$



*Since W^* is the supremum of affine functions it is convex.

Therefore,

$$\min_f \int_\Omega L(\nabla f)dx - \int_{\partial\Omega} f\cdot g\, ds = \min_f \max_\sigma \int_\Omega \langle \sigma, \nabla f\rangle - W^*(\sigma)\,dx - \int_{\partial\Omega} f g\, ds$$

$$= \min_f \max_\sigma \Big\{ \int_{\partial\Omega} ((\sigma\cdot\vec{n}) f\, ds - f g)\, ds - \int_\Omega ((\nabla\cdot\sigma) f + W^*(\sigma))\, dx.$$

If we can switch min and max. Then,

$$\min_f \int_\Omega L(\nabla f)dx - \int_{\partial\Omega} f g\, ds = \max_\sigma \min_f \int_{\partial\Omega} ((\sigma\cdot\vec{n}) f\, ds - f g)\, ds - \int_\Omega (\nabla\cdot\sigma) f + W^*(\sigma)\, dx.$$

$$\Rightarrow \boxed{\min_f \int_\Omega L(\nabla f)dx - \int_{\partial\Omega} f g\, ds = \max_{\substack{\nabla\cdot\sigma=0\\ \sigma\cdot\vec{n}=f \text{ in } \partial\Omega}} -\int_\Omega W^*(\sigma)\,dx} \to (D).$$

If $\nabla\cdot\sigma \neq 0$ or $\sigma\cdot\vec{n} \neq f$ then minimum would $= -\infty$.

Why is min max = max min?

1. $\min_y F(x,y) \leq F(x, y_0)$

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

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

2. Pointwise inequality

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

$$\Rightarrow \int_\Omega L(\nabla f) - \int_{\partial\Omega} f\cdot g\, ds \geq -\int_\Omega W^*(\sigma).$$

Equality is nontrivial in general, it is really a saddle point principle. If (P) and (D) have E-L equations then we have a direct proof:

$*$ E-L: $\nabla\cdot\left(\frac{\partial L}{\partial \nabla f}\right) = 0$ in Ω, $\frac{\partial L}{\partial \nabla f} = g$ on $\partial\Omega$

Take $\sigma = \left.\frac{\partial L}{\partial \nabla f}\right|_{f=f^*}$.

5. Example 1.

Let $\lambda_0 = $ 1st Dirichlet eigenvalue of $\Omega \subset \mathbb{R}^n$

$$\boxed{\lambda_0 = \min_{f=0 \text{ on } \partial\Omega} \frac{\int_\Omega |\nabla f|^2 dx}{\int_\Omega f^2 dx} = \min_{\substack{f=0 \text{ on } \partial\Omega \\ \int_\Omega f^2 dx = 1}} \int_\Omega |\nabla f|^2 dx.}$$

1. Sufficient to consider $f \geq 0$
2. Let $g = u^2$

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

3. The function $p^2/4z$ is convex and in fact:

$$\frac{p^2}{4z} = \max_\sigma \{\langle \sigma, p\rangle - z|\sigma|^2\}.$$ (Legendre Transform).

I already switched the order

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

$$= \max_\sigma \min_{A} \underbrace{\int_{\partial\Omega} g\,\sigma\cdot n}_{0} - \int_\Omega [\nabla\cdot\sigma + |\sigma|^2] g\, dx.$$

$$= \max_{-[\nabla\cdot\sigma + |\sigma|^2] \geq \nu} \nu$$

Again we place restrictions on σ that prevent going to $-\infty$.

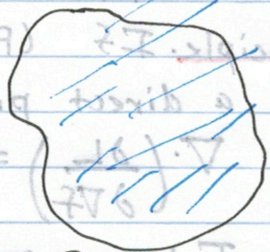
$=$ largest const ν such that $\exists$ vector field σ on Ω with $\nabla\cdot\sigma + |\sigma|^2 = -\nu$.

6. Example 2

The problem is:

$$\boxed{\min_{\nabla\cdot\sigma = 1 \text{ on } \Omega} \|\sigma\|_{L^\infty}.}$$

$\Omega \subset \mathbb{R}^2$.



Ω

It is raining on Ω, the water wants to minimize accumulation.

Observation: It is equivalent to solve

$$\boxed{\max_{\substack{|\sigma|\le 1\\ \nabla\cdot\sigma=\lambda}} \lambda}$$

Since if λ_{max} is optimal then

$$\left.\begin{array}{l}\nabla\cdot\sigma = \lambda \text{ (constant)}\\ |\sigma|\le 1\end{array}\right. \Rightarrow \lambda \le \lambda_{max}$$

$$\Rightarrow \begin{array}{l}\nabla\cdot\left(\frac{\sigma}{\lambda}\right)=1\\ |\sigma/\lambda| = \frac{1}{\lambda}\end{array} \Rightarrow \frac{1}{\lambda} \ge \frac{1}{\lambda_{max}}.$$

Therefore, $1/\lambda_{max}$ is the optimal solution for the first problem.

We now construct a dual problem

$$\boxed{\max_{\substack{|\sigma|\le 1\\ \nabla\cdot\sigma=\lambda}} \lambda = \max_{|\sigma|\le 1} \min_{\substack{f=0 \text{ on } \partial\Omega\\ \int_\Omega f dx = 1}} -\int_\Omega \langle \sigma, \nabla f\rangle dx.}$$

$$-\int_\Omega \langle \sigma, \nabla f\rangle dx = -\int_{\partial\Omega} \underbrace{(\sigma\cdot n) f}_{0} ds + \int_\Omega (\nabla\cdot\sigma) f dx = +\lambda \int_\Omega f dx = +\lambda.$$

Switching, we have that

$$\max_{\substack{|\sigma|\le 1\\ \nabla\cdot\sigma=\lambda}} \lambda = \min_{\substack{f=0 \text{ on } \partial\Omega\\ \int_\Omega f dx=1}} \max_{|\sigma|\le 1} -\int_\Omega \langle \sigma, \nabla f\rangle dx$$

$$\boxed{\Rightarrow \max_{\substack{|\sigma|\le 1\\ \nabla\cdot\sigma=\lambda}} \lambda = \min_{\substack{f=0 \text{ on } \partial\Omega\\ \int f dx=1}} \int_\Omega |\nabla f|^{?} dx}$$

Note:

$$\begin{array}{l}|\sigma|\le 1,\ \nabla\cdot\sigma = \lambda\\ f=0 \text{ on } \partial\Omega,\ \int_\Omega f dx = 1\end{array} \Longrightarrow \underbrace{-\int_\Omega \langle\sigma, \nabla f\rangle}_{\lambda} \le \int_\Omega |\nabla f| dx$$

So, $(P) \le (D)$.

This is known as an L^1-L^∞ pair. Generally, one is much easier to solve than the other.

$$* \min_{\substack{f=0 \text{ on } \partial\Omega \\ \int_\Omega f=1}} \int_\Omega |\nabla f|\,dx = \min_{D\subset\Omega} \frac{\text{length}(\partial D)}{\text{Area}(D)}. \;**$$

~~proof~~:
A key idea is the Coarea formula:

$$\int \bar{f}(x)|\nabla v|\,dx = \int\left(\int_{v=t} \bar{f}\,ds\right)dt$$

(If $v=|x-x_0|$ this gives the method of shells).

1. $$\min_{\substack{f=0 \text{ at } \partial\Omega \\ \int f\,dx=1}} \int|\nabla f|\,dx = \min_{f=0 \text{ at } \partial\Omega} \frac{\int_\Omega |\nabla f|\,dx}{\int_\Omega f\,dx}.$$

2. We can assume $f>0$ since changing f to $|f|$ leaves the |derivative| unchanged (thanks Mamikon).

3. For $f\geq 0$, $\int_\Omega f\,dx = \int_\Omega\left(\int_0^{f(x)} 1\,dt\right)dx = \int_0^\infty \text{Area}\{f\geq t\}\,dt$.
Fubini

4. $\int_\Omega |\nabla f|\,dx = \int_0^\infty \text{length}\{f=t\}\,dt$.

5. Let α be the minimum of $**$. Then,
~~$\int_0^\infty \text{length}\{f=t\}\,dt \geq \alpha\int_0^\infty \text{Area}\{f\geq t\}\,dt$.~~
$$\text{length}\{f=t\} \geq \alpha\,\text{Area}\{f\geq t\}$$
$$\Rightarrow \int_\Omega |\nabla f|\,dx \geq \alpha\int_\Omega f\,dx.$$

$$\Rightarrow * \geq **.$$

6. The opposite inequality is obtained by take a characteristic function.

<u>Cheeger's Inequality</u>: If $\lambda_0 = $ 1st Dirichlet eigenvalue of Δ in Ω and

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

Then,

$$\boxed{\frac{h^2}{4} \leq \lambda_0}$$

~~proof~~:

$\exists \sigma$ s.t. $|\sigma| \leq 1$ and $\nabla \cdot \sigma = h$. Let f_0 be the first Dirichlet Eigen function. Then,

$$h \int_\Omega f_0^2 = \int_\Omega \nabla \cdot \sigma \, f_0^2 = -2 \int_\Omega f_0 (\sigma \cdot \nabla f) \, dx$$

$$\leq 2 \int_\Omega |f_0| \cdot |\nabla f_0| \, dx$$

$$\leq 2 \|f_0\|_{L^2} \cdot \|\nabla f_0\|_{L^2}$$

$$\Rightarrow h \leq 2 \frac{\|\nabla f_0\|_{L^2}}{\|f_0\|_{L^2}} = 2 \lambda_0^{1/2}. \quad \blacksquare$$