The Lax-Milgram theorem and its applications

Brian Thompson
Advisor: Dr. David Hartenstine

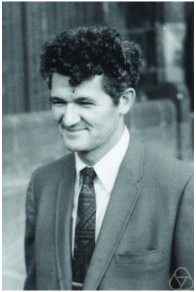
Western Washington University

Master's Project Colloquium
November 20, 2008

So, what is the Lax-Milgram theorem, and what is it good for?

- Discovered by Peter Lax and Arthur Milgram in 1954
  - (“Parabolic equations”, Contributions to the theory of partial differential equations, Annals of Mathematics Studies, no. 33. Princeton University Press, 167190.)
- Extension of the Riesz representation theorem for bilinear forms
- One application is that it shows the existence of *weak solutions* to some partial differential equations
  - In particular, we will see how it applies to the Dirichlet and Neumann problems for some elliptic partial differential equations

Personality behind the theorem: Peter D. Lax



- Came to the US from Hungary as a refugee in 1941
- Was one of the youngest people working on the Manhattan Project at age 19
- Finished his Ph.D. at NYU in 1949, only two years after his bachelor's degree.
- Best known for his work on nonlinear PDE's
- Recipient of the Abel, Wolf, and Norbert Wiener prizes, as well as the National Medal of Science.
- Single-handedly diffused a bomb set up by student radicals in the 60's at NYU

A quick review

- What is a Hilbert space?
  - An inner product space that is complete (i.e., all Cauchy sequences converge within this space) with respect to the norm / metric derived from the inner product. We will denote the i.p. of a Hilbert space with $(\cdot,\cdot)$.
- What is a a linear functional and dual space?
  - A linear functional is a linear function from a real / complex vector space to its scalar field (i.e., $\mathbb{R}$ or $\mathbb{C}$, respectively). We often denote a functional acting on an element v as $\langle f, v\rangle$.
  - A dual space is the space of all continuous linear functionals on a vector space.
- What is a bilinear form?
  - Let V be a real vector space. A mapping $B : V \times V \mapsto \mathbb{R}$ is a *bilinear form* if for any $u, u', v \in V$ and $\lambda \in \mathbb{R}$
    (i) $B[u, u' + v] = B[u, u'] + B[u, v]$,
    (ii) $B[u + v, u'] = B[u, u'] + B[v, u']$, and
    (iii) $B[\lambda u, v] = B[u, \lambda v] = \lambda B[u, v]$.
  - You can think of this as a "two-slotted linear functional"–if one slot is held constant, we have a linear functional.

•Riesz representation theorem

- Let H be a Hilbert space, and let f be in the dual of H. Then there is a unique $u \in H$ such that $(u, v) = \langle f, v \rangle$ for all v in H.
- This is what forms the basis of the Lax-Milgram theorem.

•Some notation:

- In this presentation, Ω always denotes an open subset of $\mathbb{R}^n$ with C^1 boundary.
- For sums from $i = 1$ or $i, j = 1$ to n, we will be dropping the "$\sum$", i.e., "$a_{i,j}$" means "$\sum_{i,j=1}^{n} a_{i,j}$".

What is the Lax-Milgram theorem?

The Lax-Milgram theorem states that if H is a real Hilbert space and if $B : H \times H \mapsto \mathbb{R}$ is a bilinear form that meets the following criteria:

1. B is *continuous*, i.e., $\exists \alpha > 0$ s.t. $|B[u,v]| \leq \alpha \|u\| \|v\| \, \forall u, v \in H$
2. B is *coercive*, i.e., $\exists \beta > 0$ s.t. $\beta \|u\|^2 \leq B[u,u] \, \forall u \in H$

Then for any $f \in H'$ (the dual space of H), there exists a unique $u \in H$ such that $B[u,v] = f(v)$ for all $v \in H$.

(One famous example of a a bilinear form that is continuous and coercive is the inner product–which shows us how closely related this is to the Riesz representation theorem!)

What are *weak solutions*?

•We first need to know a little bit about *classical* and *weak* solutions.

•A *classical solution* is a full-blown solution in the standard sense

- The type of solutions you would see in an introductory differential equations course; e.g., e^x is a classical solution to the differential equation $\frac{dy}{dx} = y$.

•A *weak* solution is a solution that satisfies a problem in some respect.

- Keep in mind that if a weak solution exists, that this doesn't guarantee that a classical solution exists!
- If we can find a weak solution, it is useful in that there is *maybe* a classical solution.
- Sometimes it's nice even just to have a weak solution, because it says we can *at least* satisfy our problem in *some way.*

•Some facts to consider:

- *Every* classical solution is also a weak solution.
- .... but *not every* weak solution is also a classical solution.

A short interlude into *distribution theory*

We will now need to know a little bit about the concept of the *derivative of distributions* and the *weak derivative*. For this we need some knowledge of *distributions* and *test functions*.

Definition: A *test function* is a function that is infinitely differentiable and has compact support. We denote the space of test functions by $\mathcal{D}(\Omega)$. (Where *support* is the set upon which the function is nonzero.)

Definition: A distribution is a linear functional T on $\mathcal{D}(\Omega)$ such that for every $\{\varphi_i\} \subset \mathcal{D}(\Omega)$ where $\underset{i\to\infty}{\varphi_i} = \varphi$, then $\underset{i\to\infty}{T(\varphi_i)} = T(\varphi)$.

Notice if we take any *locally integrable* function $T \in L^1_{loc}(\Omega)$, the formula $\langle T, v\rangle = \int_\Omega T(x)v(x)\,dx$ defines a distribution on Ω!

Distribution theory (cont.)

Definition: For any multiindex $\alpha \in \mathbb{N}^n$, we will set $D^\alpha = \frac{\partial^{\alpha_1 + \cdots + \alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$ (Where $\mathbb{N}$ denotes the natural numbers *with* 0).

Now we can define the *derivative of distributions*:

Definition: Let T be a distribution on $\mathcal{D}(\Omega)$. We define the α*-distributional derivative* to be: $\langle D^\alpha T, \varphi \rangle := (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle$.

Definition: Let $f \in L^1_{loc}(\Omega)$. If there is some $u \in L^1_{loc}(\Omega)$ such that $\int_\Omega f(x) D^\alpha \varphi(x) dx = (-1)^{|\alpha|} \int_\Omega u(x) \varphi(x) dx$, we call u the *weak* α*-derivative* of f.

Notice the subtlety of the difference between the *weak* and *distributional* derivatives!

We note that if T (the function, not the distribution) is in $C^1(\Omega)$, its strong derivative *is* its weak derivative. We can see this in the one-dimensional case with $\Omega = (a, b) \subset \mathbb{R}$, and extend the theory (via induction) to $\mathbb{R}^n$ and multiple partial derivatives:

$(T(x)\varphi(x))' = T'(x)\varphi(x) + T(x)\varphi'(x)$. Since the support of $T(x)\varphi(x)$ is compact and contained in (a, b) (which is open), we can see that the "edges" of $T(x)\varphi(x)$ are zero! So if we integrate over (a, b), we get:

$\int_a^b (T(x)\varphi(x))' dx = \int_a^b T'(x)\varphi(x)\,dx + \int_a^b T(x)\varphi'(x)\,dx = 0$, i.e.,

$\int_a^b T'(x)\varphi(x)dx = -\int_a^b T(x)\varphi'(x)dx$. We see that T', the proper, classical derivative of T, is also a weak derivative!

(We can now see that all classical derivatives are weak derivatives, and all weak derivatives are distributional derivatives!)

What is a Sobolev space?

We are primarily concerned with the Sobolev spaces $H^1(\Omega)$ and $H^1_0(\Omega)$. Here we let Ω be an open subset of $\mathbb{R}^n$ with a C^1 boundary, and $\alpha \in \mathbb{N}^n$. (Where $\mathbb{N}$ *includes* 0.)

Definition: $H^1(\Omega) := \{v \in L^2(\Omega) : D^\alpha v \in L^2(\Omega) \text{ s.t. } |\alpha| \leq 1\}$, with the inner product, $(u, v)_{H^1} := (u, v)_{L^2} + \sum_{i=1}^{n} (D^{x_i} u, D^{x_i} v)_{L^2}$. (Note that all of these are all weak derivatives!)

Definition: $H^1_0(\Omega)$ is the closure of $\mathcal{D}(\Omega)$ on Ω within $H^1(\Omega)$.

This is to say, this is the space of functions in $H^1(\Omega)$ that are limits of sequences of test functions. We note these facts:

- $\mathcal{D}(\Omega) \subset H^1_0(\Omega)$
- All smooth functions that are 0 on $\partial\Omega$ are contained in $H^1_0(\Omega)$.

Application 1: The Dirichlet problem

We want to find $u : \Omega \mapsto \mathbb{R}$ such that:

$$\begin{cases} -\frac{\partial}{\partial x_i}(a_{ij}\frac{\partial u}{\partial x_j}) + au = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

Where $a_{ij}, a \in L^\infty(\Omega)$ $(i, j \in \{1, ..., n\})$, where Ω is open in $\mathbb{R}^n$ with C^1 boundary, and:

1. $\mathrm{a}_{ij}(x)\xi_i\xi_j \geq \alpha|\xi|^2$ for all $\xi \in \mathbb{R}^n$, a.e. $x \in \Omega$ (for some $\alpha > 0$.) (elliptic condition)
2. $a(x) \geq 0$ a.e. $x \in \Omega$
3. Ω is bounded in one direction, or $a(x) \geq \beta > 0$ a.e. $x \in \Omega$ (for some $\beta > 0$).

Notice if we set $(a_{ij}) = Id_{n\times n}$, $a = 0$, we get the *Poisson problem:* $\Delta u = f$ on Ω, $u = 0$ on $\partial\Omega$.

Now, for some f in the dual of $H_0^1(\Omega)$, consider the problem

$$\begin{cases} u \in H_0^1(\Omega) & (i) \\ \int_\Omega a_{ij}(x)\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i} + a(x)uvdx = \langle f, v\rangle \quad \forall v \in H_0^1(\Omega) & (ii). \end{cases}$$

Notice that the integral is a bilinear form with respect to u and v. We know that $H_0^1(\Omega)$ is a real Hilbert space. It follows from the previous conditions that this bilinear form is coercive. Since a and the a_{ij} functions are bounded, we will get that the form is continuous. Thus, by Lax-Milgram, we have that there exists a unique u in $H_0^1(\Omega)$ that satisfies this problem!

... but how does this relate to the problem from our previous slide?

Let $f \in L^2(\Omega)$, and $\langle f, v \rangle = \int_\Omega f(x)v(x)dx$. Since $\mathcal{D}(\Omega) \subset H_0^1(\Omega)$, let's consider $v \in \mathcal{D}(\Omega)$.

It's easy to see that $a_{ij}D^{x_j}u, au \in L^2(\Omega)$, and thus define distributions.

We can now say that $-D^{x_i}(a_{ij}D^{x_j}u) + au = f$ in a *distributional* sense, and that since $u \in H_0^1(\Omega)$, we have that u satisfies the Dirichlet boundary condition in a weak sense as well. We thus have our weak solution!

Application 2: The Neumann problem

For a given function f, we want to find a function $u : \Omega \mapsto \mathbb{R}$ that satisfies the following Neumann problem: $\begin{cases} -\Delta u + u = f \text{ in } \Omega \\ \frac{\partial u}{\partial \nu} = 0 \text{ on } \partial\Omega \end{cases}$

(Where $\frac{\partial u}{\partial \nu}$ denotes the normal derivative of u on the boundary. Let ν denote the outward normal derivative on Ω. We define $\frac{\partial u}{\partial \nu} := \nu \cdot \nabla u$.)

Let f be a given function in $L^2(\Omega)$. We want to find some u s.t.

$$\begin{cases} u \in H^1(\Omega) & (i) \\ \int_\Omega \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i} + uv = \int_\Omega fvdx \text{ for all } v \in H^1(\Omega) & (ii) \end{cases}$$

The integral in (ii) is a bilinear form, and it follows (very similar to before) that it is continuous and coercive (Notice that this is also elliptic). Again, since $H^1(\Omega)$ is a real Hilbert space, and since the RHS integral in (ii) is of course a continuous linear functional, we have that there is a unique solution u to this problem by the L-M theorem.

Note that if we take $v \in \mathcal{D}(\Omega)$, we have that $-\frac{\partial}{\partial x_i}(\frac{\partial u}{\partial x_i}) + u = f$ in a distributional sense. This, of course, means that $-\Delta u + u = f$ in Ω in a distributional sense. We now have our *weak solution* to the first part of the Neumann problem.

Now, assume that u is C^2, and v is smooth. This assumption now gives us, in a classical sense: $\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i} = \frac{\partial}{\partial x_i}(\frac{\partial u}{\partial x_i} v) - (\frac{\partial^2 u}{\partial x_i^2})v$.

This yields the following equation: $\int_\Omega \frac{\partial}{\partial x_i}(\frac{\partial u}{\partial x_i} v) - \{\frac{\partial}{\partial x_i}(\frac{\partial u}{\partial x_i}) - u\}v = \int_\Omega fv$.

Comparing the previous equations, it is clear that $\int_\Omega \frac{\partial}{\partial x_i}(\frac{\partial u}{\partial x_i} v)dx = 0$. (Note that even though v isn't *necessarily* a test function here, we can still use the distributional derivative!)

By the divergence theorem, we have that: $\int_{\partial\Omega} \frac{\partial u}{\partial x_i} \nu_i v d\sigma(x) = 0$. (This is where the unit normal to $\partial\Omega$ is $\nu = (\nu_1, ..., \nu_n)$).

The prior integral gives us the fact: $\frac{\partial u}{\partial x_i} \nu_i = 0$, $d\sigma$ a.e. on $\partial\Omega$ for $i \in \{1, ..., n\}$. This gives us $\frac{\partial u}{\partial x_1} \nu_1 + ... + \frac{\partial u}{\partial x_n} \nu_n = \nabla u \cdot \nu = \frac{\partial u}{\partial \nu} = 0$, for $d\sigma$ a.e. on $\partial\Omega$

In sum: The Lax-Milgram theorem gives us that there exists a u that satisfies the Neumann problem weakly in two different ways: the first in a distributional sense, the second in an "almost everywhere" sense.

How do we prove the Lax-Milgram theorem?

•We start with the hypothesis that B is a continuous and coercive bilinear form on a real Hilbert space H, and that f is a continuous functional on H.

•Notice that the mapping $v \mapsto B[u_0, v]$ is a linear functional on H when we hold u_0 constant.

•By the Riesz representation theorem, $\exists$ unique Au_0 s.t. $B[u_0, v] = (Au_0, v)$ for all $v \in H$. We accordingly define the mapping $A : H \mapsto H$.

•A is linear and continuous. (Linearity comes from the inner product's linearity, continuity comes from B's continuity.)

•AH is dense in H, by B's coercivity.

•$\exists \widetilde{f} \in H$ s.t. $f(v) = (\widetilde{f}, v)$ for all $v \in H$

•By the definition of density, $\exists$ a sequence $Au_n \subset AH$ s.t. $Au_n \to \widetilde{f}$.

•u_n is bounded in H, due to B's coercivity and Cauchy-Schwarz

•$\exists$ a weakly convergent subsequence of u_n, $u_{n'} \rightharpoonup u$

•By the properties of real convergence, we see that $\langle f, v\rangle = B[u, v]$

•We can conclude that u is unique by coercivity.

Further References

If you found this presentation interesting, you ought to check out these books:

1. *Elements of Nonlinear Analysis*, Michel Chipot, Birkhäuser, 2000. This is an excellent introduction to nonlinear analysis, and has an entire chapter dedicated to the theorem and some applications. It also has some nice applications to the theorem in classical physics. This was my primary reference for my project.
2. *Partial Differential Equations*, Lawrence Evans, AMS, 1998. This is a very concise introduction to partial differential equations. There's a good introduction to Sobolev spaces here, as well as a slightly different proof of the Lax-Milgram theorem.