Tutorial: The Lax-Milgram theorem

It is a fairly simple abstract principle from linear functional analysis which provides in certain circumstances the existence and uniqueness of a weak solution to a boundary-value problem.

Let $\mathcal{H}$ be a complex Hilbert space and let $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form (which means: B is linear in both entries). Note first that:

$$B \text{ is continuous} \quad \Leftrightarrow \quad \exists M \geqslant 0 \text{ such that } |B(x,y)| \leqslant M\|x\|\,\|y\| \quad \forall x,y \in \mathcal{H}\,. \tag{1}$$

Indeed "$\Leftarrow$" follows from

$$|B(x,y) - B(x_0,y_0)| \;\leqslant\; |B(x-x_0,y) + B(x_0,y-y_0)| \;\leqslant\; \|x-x_0\|\,\|y\| + \|x_0\|\,\|y-y_0\|\,,$$

while "$\Rightarrow$" is a consequence of the uniform boundedness principle (Banach-Steinhaus), because the family $\mathcal{F} = \{B(\cdot,y)\,|\,y \in \mathcal{H}, \|y\| = 1\}$ consists by assumption of bounded linear operators $\mathcal{H} \to \mathbb{C}$ (owing to the continuity of $B(\cdot,y)$) and for each fixed $x \in \mathcal{H}$ has the property $|B(x,y)| \leqslant M_x$ $\forall B \in \mathcal{F}$ (owing to the continuity of $B(x,\cdot)$), thus by uniform boundedness $\|B(\cdot,y)\|_{\mathcal{H}^*}$ is bounded *uniformly* in y for all $B \in \mathcal{F}$ (i.e., for all y with $\|y\| = 1$), which reads precisely $|B(x,y)| \leqslant M\|x\|\,\|y\|$ $\forall x,y \in \mathcal{H}$.

In fact, the above arguments proves also that B is continuous $\Leftrightarrow$ B is separately continuous in each entry.

As a second preliminary remark, note that

$$B \text{ is continuous} \quad \Leftrightarrow \quad \exists A \in \mathcal{B}(\mathcal{H}) \text{ such that } B(x,y) = \langle x, Ay\rangle \;\forall x,y \in \mathcal{H}\,. \tag{2}$$

Indeed "$\Leftarrow$" is obvious and "$\Rightarrow$" is a consequence of the Riesz representation theorem, as follows. By assumption $B(\cdot,y)$ is a bounded linear functional $\mathcal{H} \to \mathbb{C}$ and therefore by Riesz $\exists!\, \widetilde{y} \in \mathcal{H}$ such that $B(x,y) = \langle x, \widetilde{y}\rangle$ $\forall x \in \mathcal{H}$. The map $A : \mathcal{H} \to \mathcal{H}$ defined by $Ay := \widetilde{y}$ (such a definition is well-posed) is linear because $\forall \lambda_1, \lambda_2 \in \mathbb{C}$ and $\forall y_1, y_2 \in \mathcal{H}$ one has

$$\begin{aligned} \langle x, A(\lambda_1 y_1 + \lambda_2 y_2)\rangle &= B(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 B(x,y_1) + \lambda_2 B(x,y_2) \\ &= \lambda_1 \langle x, Ay_1\rangle + \lambda_2\langle x, Ay_2\rangle \;=\; \langle x, \lambda_1 Ay_1 + \lambda_2 Ay_2\rangle\,, \end{aligned}$$

and A is bounded because $\|Ax\|^2 = \langle Ax, Ax\rangle = B(x, Ax) \leqslant M\|x\|\,\|Ax\|$, whence $\|Ax\| \leqslant M\|x\|$ $\forall x \in \mathcal{H}$.

The Lax-Milgram theorem is formulated for continuous sesquilinear forms that are also **coercive**, that is, such that $B(x,x) \geqslant c\|x\|^2$ for some $c > 0$. They exhibit the following remarkable property.

Theorem. If B is continuous and coercive on $\mathcal{H}$ then given $w \in \mathcal{H}$ there exists a unique element $x \in \mathcal{H}$ such that $B(u,x) = \langle u, w\rangle$ for all $u \in \mathcal{H}$. For such x one has $\|x\| \leqslant \frac{1}{c}\|w\|$ where $c > 0$ is the bound from below of the form (that is, $B(u,u) \geqslant c\|u\|^2$ $\forall u \in \mathcal{H}$).

Proof. Eventually $x = A^{-1}w$ where A is the bounded linear operator that represents B, see (2). Thus, all what one needs to prove is that A is invertible with bound $\|A\| \leqslant c^{-1}$, and that such x is unique. The fact that A is injective and that $\mathrm{Ran}A$ is closed in $\mathcal{H}$ follows immediately from coercivity: $c\|x\|^2 \leqslant B(x,x) = \langle x, Ax\rangle \leqslant \|x\|\,\|Ax\|$, whence $c\|x\| \leqslant \|Ax\|$ $\forall x \in \mathcal{H}$. The fact that A is surjective follows by contradiction, for otherwise $\exists z \in \mathcal{H}$, $z \neq 0$, $z \perp \mathrm{Ran}A$ (recall that $\mathrm{Ran}A$ is closed), whence $0 < c\|z\|^2 = B(z,z) = \langle z, Az\rangle = 0$. The bound $\|A\| \leqslant c^{-1}$ follows from $c\|x\| \leqslant \|Ax\|$ $\forall x \in \mathcal{H}$. Last, uniqueness holds because $B(x,y) = \langle w, y\rangle = B(\widetilde{x}, y) \Rightarrow B(x - \widetilde{x}, y) = 0$ $\forall y \in \mathcal{H}$ whence $0 \leqslant c\|x - \widetilde{x}\|^2 \leqslant B(x-\widetilde{x}, x-\widetilde{x}) = 0$ and therefore necessarily $\widetilde{x} = x$. Note that in all four steps above coercivity was used in a crucial way.

Here is a sketch of the typical application of Lax-Milgram to (elliptic) PDE*s* (you don't need to know all notions in advance, just grab the main message).

The task is always to identify the "good" Hilbert space of functions among which we look for solutions to a given PDE, and to check the validity of the assumptions on B. The sesquilinear form emerges naturally when testing the PDE against some test functions (in fact, one looks for *weak* solutions). Coercivity encodes some kind of Sobolev embedding.

Let Ω be a bounded region in $\mathbb{R}^d$. Let $f \in L^2(\Omega)$ be given. Let $a \geqslant 0$. Consider the boundary-value problem

$$\begin{cases} -\Delta u + au &= f \quad \text{in } \Omega \\ \qquad\quad u &= 0 \quad \text{on } \partial\Omega\,. \end{cases} \tag{3}$$

The claim is: there exists a unique "weak" solutions u to (3) in the space $H_0^1(\Omega)$, the completion of $C_0^\infty(\Omega)$ with respect to the norm $\|v\|_{H_0^1} := \sqrt{\|v\|_2^2 + \|\nabla v\|_2^2}$. Note that the condition $u \in H_0^1(\Omega)$ encodes the vanishing of u at the boundary of Ω. By weak solution to (3) one means a function u that satisfies

$$\langle \nabla v, \nabla u\rangle + a\langle v, u\rangle = \langle v, f\rangle \qquad \forall v \in H_0^1(\Omega)\,, \tag{4}$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in L^2 as usual. (∇u for $u \in H_0^1(\Omega)$ is a well defined function in L^2 via a limiting procedure – recall the definition of $H_0^1(\Omega)$). In fact, if u and v were smooth and vanished on $\partial\Omega$ then, owing to integration by parts, (4) would be equivalent to

$$\int_{\mathbb{R}^d} \big(-(\Delta u)(x) + au(x)\big)\, v(x)\,\mathrm{d}x \;=\; \int_{\mathbb{R}^d} f(x)\, v(x)\,\mathrm{d}x\,,$$

that is, the PDE $-\Delta u + au = f$ "tested" against v. Thus, (4) would be certainly satisfied by a "classical" solution u to (3) – which might not exist in this case, though, because f is a priori not smooth enough. (4) suggests that the appropriate bilinear form in this case is

$$B(v,u) \;:=\; \langle \nabla v, \nabla u\rangle + a\langle v, u\rangle\,. \tag{5}$$

Such a B is bounded on $H_0^1(\Omega)$ because $\forall u, v \in H_0^1(\Omega)$

$$|B(v,u)| \;\leqslant\; \|\nabla v\|_2\|\nabla u\|_2 + a\|v\|_2\|u\|_2 \;\leqslant\; (1+a)\|v\|_{H_0^1}\|u\|_{H_0^1} \tag{6}$$

(Schwartz inequality) and is coercive on $H_0^1(\Omega)$ because $\forall v \in H_0^1(\Omega)$

$$|B(v,v)| \;=\; \|\nabla v\|_2^2 + a\|v\|_2^2 \;\overset{(a\geqslant 0)}{\geqslant}\; \|\nabla v\|_2^2 \;\geqslant\; c\,\|v\|_2^2\,, \tag{7}$$

where the last step is the "deep" one and follows from the Poincaré's inequality for functions $g \in H_0^1(\Omega)$

$$\|g\|_{L^2(\Omega)} \;\leqslant\; c_\Omega \|\nabla g\|_{L^2(\Omega)} \qquad (\text{when } |\Omega| < \infty)$$

and $c := \min\{\frac{1}{2}, \frac{1}{2c_\Omega^2}\} > 0$. Therefore Lax-Milgram says that $\exists!\; u \in H_0^1(\Omega)$ such that $B(v,u) = \langle v, f\rangle$, which means precisely that there is a unique weak solution u to (3).

The original reference for the Lax-Milgram is the work: P. D. Lax, A. N. Milgram, "*Parabolic equations*" in Contributions to the theory of partial differential equations, Annals of Mathematics Studies **33** (1964) 167-190, Princeton University Press.

Alessandro