MEF I
Lecture 3 : Galerkin approximation

Pr. Ismail MERABET

Univ. K-M-Ouargla

November 3, 2024

Given a linear variational problem

$$\text{find } u \in V \text{ such that } a(u,v) = F(v) \text{ for all } v \in V,$$

we form its Galerkin approximation over a closed subspace $V_h \subset V$

$$\text{find } u_h \in V_h \text{ such that } a(u_h,v_h) = F(v_h) \text{ for all } v_h \in V_h.$$

Given a linear variational problem

$$\text{find } u \in V \text{ such that } a(u, v) = F(v) \text{ for all } v \in V,$$

we form its Galerkin approximation over a closed subspace $V_h \subset V$

$$\text{find } u_h \in V_h \text{ such that } a(u_h, v_h) = F(v_h) \text{ for all } v_h \in V_h.$$

We first consider its approximation properties over *arbitrary* subspaces V_h, then in subsequent lectures consider V_h constructed via finite elements.

Corollary

Let a and F satisfy the hypothesis of the Lax–Milgram Theorem. Then the Galerkin approximation is well-posed for any closed subspace $V_h \subset V$.

Corollary

Let a and F satisfy the hypothesis of the Lax–Milgram Theorem. Then the Galerkin approximation is well-posed for any closed subspace $V_h \subset V$.

Proof.

As $V_h \subset V$, $a : V_h \times V_h \to \mathbb{R}$ is bounded and coercive on V_h, with the same continuity and coercivity constants.

Corollary

Let a and F satisfy the hypothesis of the Lax–Milgram Theorem. Then the Galerkin approximation is well-posed for any closed subspace $V_h \subset V$.

Proof.

As $V_h \subset V$, $a : V_h \times V_h \to \mathbb{R}$ is bounded and coercive on V_h, with the same continuity and coercivity constants. $F : V_h \to \mathbb{R}$ is linear and bounded.

Corollary

Let a and F satisfy the hypothesis of the Lax–Milgram Theorem. Then the Galerkin approximation is well-posed for any closed subspace $V_h \subset V$.

Proof.

As $V_h \subset V$, $a : V_h \times V_h \to \mathbb{R}$ is bounded and coercive on V_h, with the same continuity and coercivity constants. $F : V_h \to \mathbb{R}$ is linear and bounded. Thus, by Lax–Milgram, the variational problem defining the Galerkin approximation is well-posed. $\square$

Corollary

Let a and F satisfy the hypothesis of the Lax–Milgram Theorem. Then the Galerkin approximation is well-posed for any closed subspace $V_h \subset V$.

Proof.

As $V_h \subset V$, $a : V_h \times V_h \to \mathbb{R}$ is bounded and coercive on V_h, with the same continuity and coercivity constants. $F : V_h \to \mathbb{R}$ is linear and bounded. Thus, by Lax–Milgram, the variational problem defining the Galerkin approximation is well-posed. $\square$

For coercive problems, well-posedness is inherited. *This is not true for noncoercive problems.* This makes discretising noncoercive problems much harder.

Once we choose a basis $\{\phi_i\}$ of V_h, the linear system is

$$Ax = b,$$

where

$$u_h = \sum_i x_i \phi_i, \quad b_i = F(\phi_i), \quad A_{ji} = a(\phi_i, \phi_j).$$

Once we choose a basis $\{\phi_i\}$ of V_h, the linear system is

$$Ax = b,$$

where

$$u_h = \sum_i x_i \phi_i, \quad b_i = F(\phi_i), \quad A_{ji} = a(\phi_i, \phi_j).$$

The linear system we must solve for our Galerkin approximation inherits useful properties of the underlying problem.

Once we choose a basis $\{\phi_i\}$ of V_h, the linear system is

$$Ax = b,$$

where

$$u_h = \sum_i x_i \phi_i, \quad b_i = F(\phi_i), \quad A_{ji} = a(\phi_i, \phi_j).$$

The linear system we must solve for our Galerkin approximation inherits useful properties of the underlying problem.

If a is symmetric, so is A:

$$A_{ji} = a(\phi_i, \phi_j) = a(\phi_j, \phi_i) = A_{ij}.$$

Once we choose a basis $\{\phi_i\}$ of V_h, the linear system is

$$Ax = b,$$

where

$$u_h = \sum_i x_i \phi_i, \quad b_i = F(\phi_i), \quad A_{ji} = a(\phi_i, \phi_j).$$

The linear system we must solve for our Galerkin approximation inherits useful properties of the underlying problem.

If a is symmetric, so is A:

$$A_{ji} = a(\phi_i, \phi_j) = a(\phi_j, \phi_i) = A_{ij}.$$

If a is coercive (hence positive-definite), so is A:

$$c^\top A c = a\left(\sum_i c_i \phi_i, \sum_i c_i \phi_i\right) \geq 0.$$

We know that the solution u satisfies

$$a(u, v) = F(v) \quad \text{for all } v \in V,$$

and thus in particular

$$a(u, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

We know that the solution u satisfies

$$a(u, v) = F(v) \quad \text{for all } v \in V,$$

and thus in particular

$$a(u, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

The Galerkin approximation $u_h \in V_h$ satisfies

$$a(u_h, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

We know that the solution u satisfies

$$a(u, v) = F(v) \quad \text{for all } v \in V,$$

and thus in particular

$$a(u, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

The Galerkin approximation $u_h \in V_h$ satisfies

$$a(u_h, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

Subtracting, we find

$$a(u - u_h, v_h) = 0 \quad \text{for all } v_h \in V_h.$$

We know that the solution u satisfies

$$a(u, v) = F(v) \quad \text{for all } v \in V,$$

and thus in particular

$$a(u, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

The Galerkin approximation $u_h \in V_h$ satisfies

$$a(u_h, v_h) = F(v_h) \quad \text{for all } v_h \in V_h \subset V.$$

Subtracting, we find

$$a(u - u_h, v_h) = 0 \quad \text{for all } v_h \in V_h.$$

This is called *Galerkin orthogonality*.

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Proof.

For any $v_h \in V_h$,

$$\alpha \|u - u_h\|_V^2 \leq a(u - u_h, u - u_h)$$

□

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Proof.

For any $v_h \in V_h$,

$$\begin{aligned} \alpha\|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \end{aligned}$$

□

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Proof.

For any $v_h \in V_h$,

$$\begin{aligned}
\alpha \|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\
&= a(u - u_h, u - v_h)
\end{aligned}$$

□

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Proof.

For any $v_h \in V_h$,

$$\begin{aligned} \alpha\|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\ &= a(u - u_h, u - v_h) \\ &\leq C\|u - u_h\|_V \|u - v_h\|_V. \end{aligned}$$

□

Let's assume that a is coercive and bounded, but not symmetric.

Lemma (Céa's Lemma)

The Galerkin approximation $u_h \in V_h$ to $u \in V$ is quasi-optimal, in that it satisfies

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_V.$$

Proof.

For any $v_h \in V_h$,

$$\begin{aligned}
\alpha\|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\
&= a(u - u_h, u - v_h) \\
&\leq C\|u - u_h\|_V \|u - v_h\|_V.
\end{aligned}$$

Dividing by α and minimising over $v_h \in V$, we obtain the result. □

Remark

This quasi-optimality result relates (the error in the PDE approximation) with (the approximating power of the space V_h). This decouples the error analysis from the specific PDE and turns the focus to constructing V_h with good approximation properties.

Remark

This quasi-optimality result relates (the error in the PDE approximation) with (the approximating power of the space V_h). This decouples the error analysis from the specific PDE and turns the focus to constructing V_h with good approximation properties.

This leads to the question: given $u \in V$, what is

$$\min_{v_h \in V_h} \|u - v_h\|_V?$$

In the finite element context, the answer will depend on the smoothness of u, the mesh size h, and the polynomial degree p.

Remark

This quasi-optimality result relates (the error in the PDE approximation) with (the approximating power of the space V_h). This decouples the error analysis from the specific PDE and turns the focus to constructing V_h with good approximation properties.

This leads to the question: given $u \in V$, what is

$$\min_{v_h \in V_h} \|u - v_h\|_V?$$

In the finite element context, the answer will depend on the smoothness of u, the mesh size h, and the polynomial degree p.

Remark

The ratio C/α is crucial. If $C/\alpha = 5$, things are fine. But if $C/\alpha = 1000$, our discretisation won't be very useful.

Now let's also assume that a is symmetric.

Now let's also assume that a is symmetric.

Recall that a defines a norm $\|v\|_a := \sqrt{a(v,v)}$ on V, with

$$\alpha\|v\|_V^2 \leq \|v\|_a^2 \leq C\|v\|_V^2,$$

where the continuity and coercivity constants are measured in the V norm.

Now let's also assume that a is symmetric.

Recall that a defines a norm $\|v\|_a := \sqrt{a(v,v)}$ on V, with

$$\alpha\|v\|_V^2 \leq \|v\|_a^2 \leq C\|v\|_V^2,$$

where the continuity and coercivity constants are measured in the V norm.

When we measure the continuity and coercivity constants *in the energy norm*, we get that $C = 1$ (by Cauchy–Schwarz) and $\alpha = 1$ (by definition).

Apply Céa's Lemma in the energy norm:

$$\begin{aligned} \|u - u_h\|_a &\leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_a \\ &= \min_{v_h \in V_h} \|u - v_h\|_a. \end{aligned}$$

Apply Céa's Lemma in the energy norm:

$$\begin{aligned}\|u - u_h\|_a &\leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_a \\ &= \min_{v_h \in V_h} \|u - v_h\|_a.\end{aligned}$$

Since $u_h \in V_h$, we must have equality, and thus *the error is optimal in the norm induced by the problem*:

$$\|u - u_h\|_a = \min_{v_h \in V_h} \|u - v_h\|_a.$$

Apply Céa's Lemma in the energy norm:

$$\begin{aligned}\|u - u_h\|_a &\leq \frac{C}{\alpha} \min_{v_h \in V_h} \|u - v_h\|_a \\ &= \min_{v_h \in V_h} \|u - v_h\|_a.\end{aligned}$$

Since $u_h \in V_h$, we must have equality, and thus *the error is optimal in the norm induced by the problem*:

$$\|u - u_h\|_a = \min_{v_h \in V_h} \|u - v_h\|_a.$$

The Galerkin approximation u_h is the *projection* of u onto V_h in the a-inner product!

What if we want to measure our error for a symmetric problem in the V-norm?

What if we want to measure our error for a symmetric problem in the V-norm?

Using the equivalences

$$\alpha \|v\|_V^2 \leq \|v\|_a^2 \leq C \|v\|_V^2,$$

we have

$$\|u - u_h\|_V \leq \frac{1}{\sqrt{\alpha}} \|u - u_h\|_a$$

What if we want to measure our error for a symmetric problem in the V-norm?

Using the equivalences

$$\alpha\|v\|_V^2 \le \|v\|_a^2 \le C\|v\|_V^2,$$

we have

$$\begin{aligned}\|u-u_h\|_V &\le \frac{1}{\sqrt{\alpha}}\|u-u_h\|_a \\ &= \frac{1}{\sqrt{\alpha}} \min_{v_h \in V_h} \|u-v_h\|_a\end{aligned}$$

What if we want to measure our error for a symmetric problem in the V-norm?

Using the equivalences

$$\alpha\|v\|_V^2 \le \|v\|_a^2 \le C\|v\|_V^2,$$

we have

$$\begin{aligned}\|u-u_h\|_V &\le \frac{1}{\sqrt{\alpha}}\|u-u_h\|_a \\ &= \frac{1}{\sqrt{\alpha}}\min_{v_h\in V_h}\|u-v_h\|_a \\ &\le \sqrt{\frac{C}{\alpha}}\min_{v_h\in V_h}\|u-v_h\|_V\end{aligned}$$

What if we want to measure our error for a symmetric problem in the V-norm?

Using the equivalences

$$\alpha\|v\|_V^2 \leq \|v\|_a^2 \leq C\|v\|_V^2,$$

we have

$$\begin{aligned}\|u-u_h\|_V &\leq \frac{1}{\sqrt{\alpha}}\|u-u_h\|_a \\ &= \frac{1}{\sqrt{\alpha}}\min_{v_h\in V_h}\|u-v_h\|_a \\ &\leq \sqrt{\frac{C}{\alpha}}\min_{v_h\in V_h}\|u-v_h\|_V\end{aligned}$$

so we improve the constant of quasi-optimality by a square root!