Least-Squares Finite Element Methods for Optimal Control Problems

Pavel Bochev
Computational Mathematics & Algorithms
Sandia National Laboratories
and
Max Gunzburger
Florida State University

Sixth SIAM Conference on Control and Its Applications
New Orleans, 2005
The Abstract Problem

The optimal control problems we consider consist of the following ingredients:

- **State variables** → describe the system being modeled
- **Control variables** → can be used to affect the state variables
- **State system** → PDE relating the state and control variables
- **Cost functional** → a function of the state and control variables

We restrict attention to

- **Linear, elliptic state systems**
- **Quadratic functionals**

Find state and control variables that minimize the given functional, subject to the state system being satisfied

**Hilbert spaces**

\[
\begin{align*}
\Theta & \rightarrow \text{control space} \\
\Phi & \rightarrow \text{state space} \\
\hat{\Phi} & \rightarrow \text{data space} \\
\tilde{\Phi} & \rightarrow \text{pivot space}
\end{align*}
\]

\[
\Phi \subseteq \hat{\Phi} \subseteq \tilde{\Phi} \subseteq \hat{\Phi}^* \subseteq \Phi^*
\]

\[
\langle \psi, \phi \rangle_{\Phi^*, \Phi} = \langle \psi, \phi \rangle_{\Phi^*, \hat{\Phi}} = (\psi, \phi)
\]

\[
\forall \psi \in \hat{\Phi}^* \subseteq \Phi^* \quad \forall \phi \in \Phi \subseteq \hat{\Phi}
\]
Objective Functional

\[ J(\phi, \theta) = \frac{1}{2} a_1(\phi - \hat{\phi}, \phi - \hat{\phi}) + \frac{1}{2} a_2(\theta, \theta) \]

\( a_1(\cdot, \cdot) \leftarrow \text{symmetric bilinear on} \ \hat{\Phi} \times \hat{\Phi} \)

\( a_2(\cdot, \cdot) \leftarrow \text{symmetric bilinear on} \ \Theta \times \Theta \)

Assumptions

\[
\begin{align*}
    a_1(\phi, \mu) &\leq C_1 \|\phi\|_{\hat{\Phi}} \|\mu\|_{\hat{\Phi}} & \leftarrow & \text{continuity} \\
    a_2(\theta, \nu) &\leq C_2 \|\theta\|_{\Theta} \|\nu\|_{\Theta} & \leftarrow & \text{continuity} \\
    a_1(\phi, \phi) &\geq 0 & \leftarrow & \text{non-negativity} \\
    a_2(\theta, \theta) &\geq K_2 \|\theta\|_{\Theta} & \leftarrow & \text{coercivity}
\end{align*}
\]

- Often \( \Theta \) is finite dimensional. Then \( \theta \) is referred to as the vector of design variables
- The second term is penalty that limits the size of the control \( \theta \)
Linear Constraint Equation

\[ b_1(\phi, \psi) + b_2(\theta, \psi) = \langle g, \psi \rangle_{\Lambda^*, \Lambda} \]

Assumptions

\[
\begin{align*}
    b_1(\phi, \psi) &\leq c_1 \|\phi\|_{\Phi} \|\psi\|_{\Lambda} & \leftarrow \text{continuity} \\
    b_2(\theta, \psi) &\leq c_2 \|\theta\|_{\Theta} \|\psi\|_{\Phi} & \leftarrow \text{continuity} \\
    \sup_{\psi \in \Lambda} \frac{b_1(\phi, \psi)}{\|\psi\|_{\Lambda}} &\geq k_1 \|\phi\|_{\Phi} & \leftarrow \text{weak coercivity} \\
    \sup_{\phi \in \Phi} \frac{b_1(\phi, \psi)}{\|\phi\|_{\Phi}} &> 0 & \leftarrow \text{weak coercivity}
\end{align*}
\]

These assumptions are sufficient to guarantee that given any control \( \theta \) the constraint equation is uniquely solvable for the state \( \phi \).
The Optimal Control Problem

\[
\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad b_1(\phi, \psi) + b_2(\theta, \psi) = \langle g, \psi \rangle_{\Lambda^*, \Lambda^*}
\]

Operator notation

\[
\begin{align*}
a_1(\cdot, \cdot) &\rightarrow A_1 : \hat{\Phi} \mapsto \hat{\Phi}^* \\
a_2(\cdot, \cdot) &\rightarrow A_2 : \Theta \mapsto \Theta^* \\
b_1(\cdot, \cdot) &\rightarrow B_1 : \Phi \mapsto \Lambda^* \\
b_2(\cdot, \cdot) &\rightarrow B_2 : \Lambda \mapsto \Theta^*
\end{align*}
\]

\[
J(\phi, \theta) = \frac{1}{2} \left\langle A_1(\phi - \hat{\phi}), (\phi - \hat{\phi}) \right\rangle_{\Phi^*, \Phi^*} + \frac{1}{2} \left\langle A_2 \theta, \theta \right\rangle_{\Theta^*, \Theta^*}
\]

\[
B_1 \phi + B_2 \theta = g
\]

\[
\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad B_1 \phi + B_2 \theta = g \quad \text{in} \quad \Lambda^*
\]
Lagrange Multiplier Solution

Saddle point optimality system

\[
\begin{align*}
& a_1(\phi, \mu) + b_1(\mu, \lambda) = a_1(\hat{\phi}, \mu) \quad \forall \mu \in \Phi
\end{align*}
\]

\[
\begin{align*}
& a_2(\theta, \nu) + b_2(\nu, \lambda) = 0 \quad \forall \nu \in \Theta
\end{align*}
\]

\[
\begin{align*}
& b_1(\phi, \psi) + b_2(\theta, \psi) = \langle g, \psi \rangle_{\Lambda^*, \Lambda} \quad \forall \psi \in \Lambda
\end{align*}
\]

\[
\begin{align*}
& A_1\phi + B_1^*\lambda = A_1\hat{\phi} \quad \text{in} \quad \Phi^*
\end{align*}
\]

\[
\begin{align*}
& A_2\theta + B_2^*\lambda = 0 \quad \text{in} \quad \Theta^*
\end{align*}
\]

\[
\begin{align*}
& B_1\phi + B_2\theta = g \quad \text{in} \quad \Lambda^*
\end{align*}
\]

Assumptions imply that (using the Brezzi theory)

– The optimal control problem has a unique solution \((\phi, \theta) \in \Phi \times \Theta\)
– That solution can be determined by solving the saddle point optimality system
– The optimality system has a unique solution \((\phi, \theta, \lambda) \in \Phi \times \Theta \times \Lambda\)
– That solution depends continuously on the data

\[
\|\phi\|_\Phi + \|\theta\|_\Theta + \|\lambda\|_\Lambda \leq C(\|g\|_{\Lambda^*} + \|\hat{\phi}\|_\Phi)
\]
Galerkin approximation of the optimality system

Approximation spaces \( \Phi^h \subset \Phi \quad \Theta^h \subset \Theta \quad \Lambda^h \subset \Lambda \)

Discrete problem

\[
\begin{align*}
& a_1(\phi^h, \mu^h) + b_1(\mu^h, \lambda^h) = a_1(\hat{\phi}^h, \mu^h) \quad \forall \mu^h \in \Phi^h \\
& a_2(\theta^h, \nu^h) + b_2(\nu^h, \lambda^h) = 0 \quad \forall \nu^h \in \Theta^h \\
& b_1(\phi^h, \psi^h) + b_2(\theta^h, \psi^h) = \langle g, \psi \rangle_{\Lambda^h, \Lambda} \quad \forall \psi^h \in \Lambda^h
\end{align*}
\]

Weak coercivity **not inherited** on subspaces \( \Rightarrow \) Discrete inf-sup conditions

\[
\begin{align*}
\sup_{\psi^h \in \Lambda^h} \frac{b_1(\phi^h, \psi^h)}{\|\psi^h\|_\Lambda} & \geq k_i^h \|\phi^h\|_\Phi \quad \text{and} \quad \sup_{\phi^h \in \Phi^h} \frac{b_1(\phi^h, \psi^h)}{\|\phi^h\|_\Phi} > 0 \\
\|\phi^h\|_\Phi + \|\theta^h\|_\Theta + \|\lambda^h\|_\Lambda & \leq C\left(\|g\|_\Lambda + \|\hat{\phi}\|_\Phi\right) \\
\|\phi - \phi^h\|_\Phi + \|\theta - \theta^h\|_\Theta + \|\lambda - \lambda^h\|_\Lambda & \leq C \inf\left(\|\phi - \mu^h\|_\Phi + \|\theta - \zeta^h\|_\Theta + \|\lambda - \psi^h\|_\Lambda\right)
\end{align*}
\]
The Algebraic Problem

The discrete optimality system is equivalent to the linear system

\[
\begin{pmatrix}
A_1 & B_1^T \\
A_2 & B_2^T \\
B_1 & B_2
\end{pmatrix}
\begin{pmatrix}
\phi \\
\theta \\
\lambda
\end{pmatrix} =
\begin{pmatrix}
f \\
0 \\
g
\end{pmatrix}
\]

The saddle point nature of optimality system cannot be avoided
– It occurs even when the state system is strongly coercive, i.e.,
   even if the form \( b_1(.,..) \) is coercive

The Galerkin approach will always yield indefinite matrix problems

The discretized optimality system is formidable
– At least twice the size of the state system
– One shot solution is often impractical
– Many strategies for uncoupling the equations have been proposed

Necessitates an iterative approach to solving the system
Least-Squares Methods I: Application to the Optimality System

Starting point

\begin{align*}
A_1 \phi + B_1^* \lambda &= A_1 \hat{\phi} \quad \text{in } \Phi^* \\
A_2 \theta + B_2^* \lambda &= 0 \quad \text{in } \Theta^* \quad \text{Well-posed in } \Phi^* \times \Theta^* \times \Lambda^* \quad \leftarrow \text{solution space} \\
B_1 \phi + B_2 \theta &= g \quad \text{in } \Lambda^* \quad \leftarrow \text{data space}
\end{align*}

Least-squares functional

\[
K(\phi, \theta, \lambda; \hat{\phi}, g) = \frac{1}{2} \left( \| A_1 \phi + B_1^* \lambda - A_1 \hat{\phi} \|_{\Phi^*}^2 + \| A_2 \theta + B_2^* \lambda \|_{\Theta^*}^2 + \| B_1 \phi + B_2 \theta - g \|_{\Lambda^*}^2 \right)
\]

Minimization problem

\[
\min_{\phi, \theta, \lambda} K(\phi, \theta, \lambda; \hat{\phi}, g)
\]
Optimality condition  

\[ B(\{\phi, \theta, \lambda\}, \{\mu, \nu, \psi\}) = F(\{\mu, \nu, \psi\}, \{A_1 \phi, 0, g\}) \]

\[ B(\{\phi, \theta, \lambda\}, \{\mu, \nu, \psi\}) = (A_1 \phi + B_1^* \lambda, A_1 \mu + B_1^* \psi)_{\phi^*} + (A_2 \theta + B_2^* \lambda, A_2 \nu + B_2^* \psi)_{\theta^*} + (B_1 \phi + B_2 \theta, B_1 \mu + B_2 \nu)_{\lambda^*} \]

\[ F(\{\mu, \nu, \psi\}, \{A_1 \phi, 0, g\}) = (A_1 \phi, A_1 \mu + B_1^* \psi)_{\phi^*} + (g, B_1 \mu + B_2 \nu)_{\lambda^*} \]

Our assumptions imply that

- \( B(., .) \) is symmetric, continuous and coercive: 
  \[ B(\{\phi, \theta, \lambda\}, \{\phi, \theta, \lambda\}) \geq C (\|\phi\|_{\phi}^2 + \|\theta\|_{\theta}^2 + \|\lambda\|_{\lambda}^2) \]

- \( F(.) \) is continuous

And since \( K(\phi, \theta, \lambda; 0, 0) = B(\{\phi, \theta, \lambda\}, \{\phi, \theta, \lambda\}) \) the least-squares functional is norm-equivalent

\[ C_1 (\|\phi\|_{\phi}^2 + \|\theta\|_{\theta}^2 + \|\lambda\|_{\lambda}^2) \leq K(\phi, \theta, \lambda; 0, 0) \leq C_2 (\|\phi\|_{\phi}^2 + \|\theta\|_{\theta}^2 + \|\lambda\|_{\lambda}^2) \]

The least-squares minimization problem has a unique minimizer and

\[ \|\phi\|_{\phi} + \|\theta\|_{\theta} + \|\lambda\|_{\lambda} \leq C (\|g\|_{\lambda^*} + \|\phi\|_{\phi^*}) \]
Least-Squares Finite Element Method

Approximation spaces

\[ \Phi^h \subset \Phi \quad \Theta^h \subset \Theta \quad \Lambda^h \subset \Lambda \]

Discrete problem

\[ B\left(\{\phi^h, \theta^h, \lambda^h\}, \{\mu^h, \nu^h, \psi^h\}\right) = F\left(\{\mu^h, \nu^h, \psi^h\}, \{A_1 \hat{\phi}, 0, g\}\right) \]

Coercivity is inherited on subspaces \(\Rightarrow\) no inf-sup conditions are needed!

\[ B\left(\{\phi^h, \theta^h, \lambda^h\}, \{\phi^h, \theta^h, \lambda^h\}\right) \geq C\left(\|\phi^h\|_{\Phi}^2 + \|\theta^h\|_{\Theta}^2 + \|\lambda^h\|_{\Lambda}^2\right) \]

Practicality of this approach requires one to cast the constraint equations as first-order systems
The Algebraic Problem

The discrete least-squares optimality system is equivalent to a linear system

\[
\begin{pmatrix}
K_1 & C_1^T & C_2^T \\
C_1 & K_2 & C_3^T \\
C_2 & C_3 & K_3
\end{pmatrix}
\begin{pmatrix}
\phi \\
\theta \\
\lambda
\end{pmatrix}
=
\begin{pmatrix}
f \\
h \\
g
\end{pmatrix}
\]

- The coefficient matrix of this system is **symmetric and positive definite**
- this should be compared to the Galerkin linear system for which the coefficient matrix is **symmetric and indefinite**
- This system is also formidable and calls for uncoupling strategies for its solution

Uncoupling strategies

\[
\begin{pmatrix}
A_1 & B_1^T \\
A_2 & B_2^T \\
B_1 & B_2
\end{pmatrix}
\begin{pmatrix}
\phi \\
\theta \\
\lambda
\end{pmatrix}
=
\begin{pmatrix}
f \\
0 \\
g
\end{pmatrix}
\]

- Rely on invertibility of \(B_1\) and \(A_2\)
- \(B_1\) is, in general, non-symmetric and indefinite

\[
\begin{pmatrix}
K_1 & C_1^T & C_2^T \\
C_1 & K_2 & C_3^T \\
C_2 & C_3 & K_3
\end{pmatrix}
\begin{pmatrix}
\phi \\
\theta \\
\lambda
\end{pmatrix}
=
\begin{pmatrix}
f \\
h \\
g
\end{pmatrix}
\]

- Rely on invertibility of \(K_1, K_2\) and \(K_2\)
- all are symmetric and positive definite
- true even if discrete inf-sup for \(b_1\) does not hold
A Simple Uncoupling Strategy

An example of a simple uncoupling strategy is to apply block-Gauss Seidel method

Start with initial guesses $\phi^{(0)}$ and $\theta^{(0)}$ for the discretized state and control; then, for $k=1,2\ldots$ successively solve the linear systems

$$
K_3\lambda^{(k+1)} = g - C_2\phi^{(k)} - C_3\theta^{(k)}
$$

$$
K_1\phi^{(k+1)} = f - C_1^T\theta^{(k)} - C_2^T\lambda^{(k+1)}
$$

$$
K_2\theta^{(k+1)} = h - C_1\phi^{(k+1)} - C_3^T\lambda^{(k+1)}
$$

until satisfactory convergence is achieved

– The matrices $K_1$, $K_2$ and $K_3$ are all symmetric positive definite so that efficient solution strategies are available

– of course, more sophisticated uncoupling strategies can also be defined
Least-Squares Methods II: Bi-level Minimization Problem

Least-squares form of the constraint

\[ B_1\phi + B_2\theta = g \]

Treat the control \( \theta \) as being a given function

Define the least-squares functional

\[ H(\phi;\theta,g) = \frac{1}{2}\|B_1\phi + B_2\theta - g\|_A^2. \]

Consider the minimization problem

\[ \min_{\phi} H(\phi;\theta,g) \]

Euler-Lagrange equation

\[ \tilde{b}_1(\phi,\mu) = (g,B_1\mu)_A - \tilde{b}_2(\theta,\mu) \quad \forall \mu \in \Phi \]

\[ \tilde{b}_2(\theta,\nu) = (B_2\theta,B_2\nu)_A. \]

Norm-equivalence:

\[ \tilde{b}_1(\phi,\phi) \geq C\|\phi\|_\Phi^2 \quad \Rightarrow \quad C_2\|\phi\|_\Phi^2 \leq H(\phi;0,0) \leq C_2\|\phi\|_\Phi^2 \]

– The least-squares form of the constraint is a well-posed problem for all \( \theta \)
– Its approximation does not require inf-sup conditions
– The discrete problem leads to symmetric and positive definite systems
Bi-level Minimization problem

Instead of considering the optimal control problem

$$\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad B_1 \phi + B_2 \theta = g \quad \text{in} \quad \Lambda^*$$

we consider the equivalent problem

$$\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad \min_{\phi} H(\phi; \theta, g)$$

– This is a bi-level optimization problem
– There are, of course, many ways to address such problems
– We will describe two possible approaches
Form the penalized functional

\[ J_\varepsilon (\phi, \theta) = J(\phi, \theta) + \frac{1}{\varepsilon} H(\phi; \theta, g) \]

Unfortunately, this approach does not take full advantage of the least-squares formulation of the constraint equation:

– in particular, the discrete inf-sup condition on the bilinear form \( b_1(\ldots) \) cannot be circumvented;

– one also has to worry about choosing a “good” value for the penalty parameter \( \varepsilon \).
Constraining by Least-Squares
First Order Necessary Conditions

The bi-level optimization problem is

\[
\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad \min_{\phi} H(\phi; \theta, g)
\]

And the first-order necessary condition for the least-squares principle in operator form is:

\[
B_1^* B_1 \phi + B_1^* B_2 \theta = B_1^* g \quad \text{in} \quad \Phi^*
\]

Therefore,

\[
\min_{\phi, \theta} J(\phi, \theta) \quad \text{subject to} \quad B_1^* B_1 \phi + B_1^* B_2 \theta = B_1^* g \quad \text{in} \quad \Phi^*
\]

is an equivalent reformulation of our optimal control problem.
The Euler-Lagrange equations corresponding to the reformulated bi-level optimization

\[ \begin{align*}
A_1 \phi + B_1^* B_1 \lambda &= A_1 \hat{\phi} \quad \text{in} \quad \Phi^* \\
A_2 \theta + B_2^* B_1 \lambda &= 0 \quad \text{in} \quad \Theta^* \\
B_1^* B_1 \phi + B_1^* B_2 \theta &= B_1^* g \quad \text{in} \quad \Lambda^*
\end{align*} \]

– This is a saddle-point problem and so the resulting matrix problem is indefinite
– The only advantage over Galerkin is that $B_1^* B_1$ is positive definite even when $B_1$ is not

There are several effective means for discretizing the new optimality system

– Penalty methods are one approach
– One also has the choice of discretize-then-optimize and optimize-then-discretize
A Taxonomy of Optimization Approaches

1. LM FE
2. LM PEN FE ELIM
3. LM LS FE
4. LS-PEN OPTIM FE
5. LS-CON LM FE
6. LS-CON LM PEN FE ELIM
7. LS-CON PEN OPTIM FE
From the table we see that only Method 6 has all its boxes checked ⇒ it seems to be the preferred method.
But…

There are additional issues that arise in practice and must also be considered:

– Use standard finite elements
– Ease of assembly of the discrete system
– Manageable conditioning of the discrete system

When these are added to the mix, it seems that method 3 wins out:

LM  LS  FE

For further details see:

