OPTIMAL DESIGN OF STRUCTURES (MAP 562)

G. ALLAIRE

January 23rd, 2019

Department of Applied Mathematics, Ecole Polytechnique

CHAPTER IV

OPTIMAL CONTROL

Optimization of distributed systems:

Computing a gradient by the adjoint method

Control of an elastic membrane

For $f \in L^2(\Omega)$, the vertical displacement u of the membrane is solution of

$$\begin{cases} -\Delta u = f + v & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where v is a **control force** which is our optimization variable (for example, a piezzo-electric actuator). We define the set of admissible controls

$$K = \left\{ v \in L^2(\omega) \mid v_{min}(x) \leq v(x) \leq v_{max}(x) \text{ in } \omega \text{ and } v = 0 \text{ in } \Omega \setminus \omega \right\}.$$

We want to control the membrane in order to reach a prescribed displacement $u_0 \in L^2(\Omega)$ by minimizing ($c > 0$)

$$\inf_{v \in K} \left\{ J(v) = \frac{1}{2} \int_\Omega \left(|u - u_0|^2 + c|v|^2 \right) dx \right\}.$$

Existence of an optimal control

Proposition.

There exists a unique optimal control $\overline{v} \in K$.

Proof. $v \to u$ is an affine function from K into $H_0^1(\Omega)$.

The integrand of J is a positive "polynomial" of degree two in v.

$v \to J(v)$ is strongly convex on K which is convex.

Remark. The existence is often more delicate to prove, but the important thing here is to compute a gradient $J'(v)$ for numerical purposes.

Important notice: the solution u of the p.d.e. depends on the control v.

Gradient and optimality condition

The safest and simplest way of computing a gradient is to evaluate the directional derivative

$$j(t) = J(v + tw) \quad \Rightarrow \quad j'(0) = \langle J'(v), w\rangle = \int_\Omega J'(v) w \, dx \; .$$

By linearity, we have $u(v + tw) = u(v) + t\tilde{u}(w)$ with

$$\begin{cases} -\Delta \tilde{u}(w) = w & \text{in } \Omega \\ \tilde{u}(w) = 0 & \text{on } \partial\Omega. \end{cases}$$

In other words, $\tilde{u}(w) = \langle u'(v), w\rangle$.

Since $J(v)$ is quadratic the computation is very simple and we obtain

$$\int_\Omega J'(v) w \, dx = \int_\Omega \Big((u(v) - u_0)\tilde{u}(w) + cvw \Big) dx,$$

Unfortunately $J'(v)$ is not explicit because we cannot factorize out w in $\tilde{u}(w)$!

Adjoint state

To simplify the gradient formula we use the so-called adjoint state p, defined as the unique solution in $H_0^1(\Omega)$ of

$$\begin{cases} -\Delta p = u - u_0 & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega. \end{cases}$$

We multiply the equation for $\tilde{u}(w)$ by p and conversely

$$\text{equation for } p \times \tilde{u}(w) \quad \Rightarrow \quad \int_\Omega \nabla p \cdot \nabla \tilde{u}(w)\, dx = \int_\Omega (u - u_0)\tilde{u}(w)\, dx$$

$$\text{equation for } \tilde{u}(w) \times p \quad \Rightarrow \quad \int_\Omega \nabla \tilde{u}(w) \cdot \nabla p\, dx = \int_\Omega wp\, dx$$

Comparing these two equalities we deduce that

$$\int_\Omega (u - u_0)\tilde{u}(w)\, dx = \int_\Omega wp\, dx \quad \Rightarrow \quad \int_\Omega J'(v)w\, dx = \int_\Omega (p + cv)w\, dx.$$

Conclusion on the adjoint state

We found an explicit formula of the gradient

$$J'(v) = p + cv.$$

- Adjoint method: computation of the gradient by solving 2 boundary value problems (u and p).
- If one does not use the adjoint: for **each** direction w one must solve 2 boundary value problems (u and $\tilde{u}(w)$) to evaluate $\langle J'(v), w\rangle$.

  For example, if $J'(v)$ is a vector of dimension n, its n components are obtained by solving $(n+1)$ problems !
- Very efficient in practice: it is the best possible method.
- Inconvenient: if one uses a black-box software to compute u, it can be very difficult to modify it in order to get the adjoint state p.

Further remarks on the notion of adjoint state

- If the state equation is not self-adjoint (the bilinear form is not symmetric), the operator of the adjoint equation is the transposed or adjoint of the direct operator.
- If the state equation is time dependent with an initial condition, then the adjoint equation is time dependent too, but backward with a final condition.
- If the state equation is non-linear, the adjoint equation is linear.

The adjoint is not just a trick ! It can be deduced from the Lagrangian of the problem.

General method to find the adjoint equation

We consider the state equation as a constraint and, for any $(\hat{v}, \hat{u}, \hat{p}) \in L^2(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega)$, we introduce the Lagrangian of the minimization problem

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega \hat{p}(\Delta \hat{u} + f + \hat{v})\, dx,$$

where $\hat{p}$ is the Lagrange multiplier for the constraint which links the two independent variables $\hat{v}$ and $\hat{u}$.

Integrating by parts yields

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega (-\nabla \hat{p} \cdot \nabla \hat{u} + f\hat{p} + \hat{v}\hat{p})\, dx.$$

Proposition. The optimality conditions are equivalent to the stationnarity of the Lagrangian, i.e.,

$$\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0.$$

$\boxed{\text{Proof}}$

- $\frac{\partial \mathcal{L}}{\partial p} = 0 \Rightarrow$ by definition, we recover the equation satisfied by the state u.
- $\frac{\partial \mathcal{L}}{\partial u} = 0 \Rightarrow$ equation satisfied by the adjoint state p. Indeed,

$$\ell_u(t) = \mathcal{L}(\hat{v}, \hat{u} + t\phi, \hat{p}) \quad \Rightarrow \quad \ell_u'(0) = \langle \frac{\partial \mathcal{L}}{\partial u}, \phi \rangle = \int_\Omega \left((\hat{u} - u_0)\phi - \nabla \hat{p} \cdot \nabla \phi \right) dx$$

which is the variational formulation of the adjoint equation.

- $\frac{\partial \mathcal{L}}{\partial v} = 0 \Rightarrow$ formula for $J'(v)$. Indeed,

$$\ell_v(t) = \mathcal{L}(\hat{v} + tw, \hat{u}, \hat{p}) \quad \Rightarrow \quad \ell_v'(0) = \langle \frac{\partial \mathcal{L}}{\partial v}, w \rangle = \int_\Omega (c\hat{v} + \hat{p}) w \, dx$$

Simple formula for the derivative

In the preceding proof we obtained

$$\boxed{J'(v) = \frac{\partial \mathcal{L}}{\partial v}(v, u, p)}$$

with the state u and the adjoint p (both depending on v).

It is not a surprise ! Indeed,

$$J(v) = \mathcal{L}(v, u, \hat{p}) \quad \forall \hat{p}$$

because u is the state. Thus, if $u(v)$ is differentiable, we get

$$\langle J'(v), w\rangle = \langle \frac{\partial \mathcal{L}}{\partial v}(v, u, \hat{p}), w\rangle + \langle \frac{\partial \mathcal{L}}{\partial u}(v, u, \hat{p}), \frac{\partial u}{\partial v}(w)\rangle$$

We then take $\hat{p} = p$, the adjoint, to obtain

$$\langle J'(v), w\rangle = \langle \frac{\partial \mathcal{L}}{\partial v}(v, u, p), w\rangle$$

Another interpretation of the adjoint state

The adjoint state p is the Lagrange multiplier for the constraint of the state equation. But it is also a sensitivity function.

Define the Lagrangian

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}, f) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega (-\nabla \hat{p} \cdot \nabla \hat{u} + f\hat{p} + \hat{v}\hat{p})\, dx.$$

We study the sensitivity of the minimum with respect to variations of f.

We denote by $v(f)$, $u(f)$ and $p(f)$ the optimal values, depending on f. We assume that they are differentiable with respect to f. Then

$$\nabla_f \Big(J(v(f)) \Big) = p(f).$$

p gives the derivative (without further computation) of the minimun with respect to f !

Indeed $J(v(f)) = \mathcal{L}(v(f), u(f), p(f), f)$ and $\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0$.