IC-32 | # Optimal Control | **Winter 2006**

Benoît Chachuat
ME C2 401, Ph: 33844, benoit.chachuat@epfl.ch

1. What is the optimal solution of an unconstrained minimization problem with linear cost function, $f(\mathbf{x}) = \mathbf{c}^\mathsf{T}\mathbf{x}$, with $\mathbf{c} \neq \mathbf{0}$?

---

2. Given a cardboard of area A to make a rectangular box, what is the maximum volume that can be attained?

---

3. Consider P points $(x_1, y_1), \ldots, (x_P, y_P) \in \mathbb{R}^2$. Given a set of basis functions, $b_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, find the coefficients $p_1, \ldots, p_n$ in such a way that the curve

$$y = \sum_{i=1}^{n} p_i b_i(x),$$

gives the minimum least square error

$$f(\mathbf{p}) := \sum_{k=1}^{P} \left(y_k - \sum_{i=1}^{n} p_i b_i(x_k) \right)^2.$$

---

4. In this problem, we consider a discrete-time dynamic system of the form

$$\begin{aligned} x_{i+1} &= f(x_i, u_i) \\ x_0 &= a \text{ (given)}, \end{aligned} \tag{1}$$

and try to find a control sequence $\mathbf{u} = (u_0, \ldots, u_{N-1})^\mathsf{T}$ that minimizes the function

$$\mathcal{J}(\mathbf{u}) \ = \ g_N(x_N) + \sum_{i=0}^{N-1} g_i(x_i, u_i).$$

In this objective, we wish to calculate the gradient of the objective function $\mathcal{J}$ with respect to the control variables $\mathbf{u}$.

I) Sensitivity-Based Approach for Gradient Calculation.

(a) Assume that a solution to the difference equation (1) exists and is unique, and can be written as an explicit function of the control variables,

$$x_i \ = \ \phi_i(u_0, \ldots, u_{i-1}).$$

Show that the gradient of the objective function $\mathcal{J}$ is given by

$$\nabla_{u_k}\mathcal{J} \;=\; \frac{\partial g_k}{\partial u_k} + \sum_{i=k+1}^{N} \frac{\partial g_i}{\partial x_i}\frac{\partial \phi_i}{\partial u_k}, \tag{2}$$

for each $k = 1, \ldots, N-1$.

(b) Using a chain rule of differentiation, derive a recursive expression for the state sensitivities, i.e., a mathematical expression giving $\frac{\partial x_{i+1}}{\partial u_k}$ as a function of $\frac{\partial x_i}{\partial u_k}$. Also derive initial conditions for the state sensitivities $\frac{\partial x_0}{\partial u_k}$.

(c) Evaluating the gradient $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{J}$ by directly substituting the state sensitivities $\frac{\partial x_i}{\partial u_k}$ in Eq. (2) is often referred to as the *(forward) sensitivity approach* in the literature. How many finite difference equations is it necessary to solve for evaluating the gradient of the objective function in this approach? And how can we expect the (computational) cost of a gradient evaluation to grow with the number N of control variables?

II) Adjoint-Based Approach for Gradient Calculation.

(a) Consider the augmented function

$$\widehat{\mathcal{J}} \;:=\; \mathcal{J} + \sum_{i=0}^{N-1} p_{i+1}\left[f_i(x_i,u_i) - x_{i+1}\right],$$

which is obviously equal to $\mathcal{J}$, as long as the state variables x_i satisfy the difference equation (1), irrespective of the values of the new variables $p_1, \ldots, p_N$. Using a chain rule of differentiation (once again), prove that the gradient of $\widehat{\mathcal{J}}$ is given by

$$\boldsymbol{\nabla}_{\mathbf{u}}\widehat{\mathcal{J}} \;=\; \sum_{i=0}^{N-1}\left[\frac{\partial g_i}{\partial \mathbf{u}} + p_{i+1}\frac{\partial f_i}{\partial \mathbf{u}}\right] + \sum_{i=1}^{N-1}\left[\frac{\partial g_i}{\partial x_i} + p_{i+1}\frac{\partial f_i}{\partial x_i} - p_i\right]\frac{\partial x_i}{\partial \mathbf{u}} + \left[\frac{\partial g_N}{\partial x_N} - p_N\right]\frac{\partial x_N}{\partial \mathbf{u}}.$$

Conclude that, for some suitable choice of the *adjoint variables* p_i (the so-called *adjoint equations*), the gradient of the objective function $\mathcal{J}$ can be calculated as

$$\nabla_{u_k}\mathcal{J} \;=\; \frac{\partial g_k}{\partial u_k} + p_{k+1}\frac{\partial f_k}{\partial u_k}, \tag{3}$$

for each $k = 1, \ldots, N-1$.

(b) Evaluating the gradient $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{J}$ from Eq. (3) is often referred to as the *adjoint* or *reverse sensitivity approach* in the literature. Propose an algorithm that calculates $\nabla_{\mathbf{u}}\mathcal{J}$ based on this approach. How many finite difference equations is it necessary to solve for evaluating the gradient of the objective function? And how can we expect the (computational) cost of a gradient evaluation to grow with the number N of control variables? Conclude.

III) Application to a Reservoir Regulation Problem.

Let x_i denote the volume of water held in a reservoir at the ith of N time periods. The volume x_i evolves according to

$$x_{i+1} \;=\; x_i - u_i, \quad i = 0, \ldots, N-1; \quad x_0 = 2,$$

where u_i is water used for some productive purpose in period i. The volume x_i can be viewed as the state, and the outflow u_i as the control.

For this problem, we define the objective function to be minimized as

$$\mathcal{J}(\mathbf{u}) \;=\; (x_N - 1)^2 + \sum_{i=0}^{N-1} \left[(x_i - 1)^2 - u_i\right].$$

That is, we want to find the control $\mathbf{u}$ that keeps the volume in the reservoir close to 1 in each time period, and at the same time, maximize the outflow u_i of water (e.g., for electric power generation). No restriction is placed on the volume x_i in the reservoir, nor on the outflow u_i.

(a) Find a mathematical expression giving the state x_i as a function of the outflows $u_0, \ldots, u_{i-1}$ in previous time periods. Conclude that the objective function is convex in $\mathbb{R}^N$.

(b) Derive the difference equation and the associated terminal condition satisfied by the adjoint variables p_k, $k = 0, \ldots, N$, and use it to obtain an expression of the gradient $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{J}$. Conclude that stationary points for this problem are those satisfying the conditions

$$p_k^\star \;=\; -1, \quad \forall k = 1, \ldots, N.$$

(c) Show that the foregoing stationarity conditions impose

$$x_i^\star = 1, \quad i = 1, \ldots, N-1, \quad \text{and} \quad x_N^\star = \frac{1}{2}.$$

Conclude that the *unique* optimal control sequence for that problem is

$$\mathbf{u}^\star = \left(1, 0, \ldots, 0, \frac{1}{2}\right)^{\mathsf{T}}.$$