Chapter 1
What is Optimal Control Theory?

Suresh P. Sethi

The University Of Texas at Dallas

June 2019
What is Optimal Control Theory?

- Dynamic Systems: Evolving over time.
- Time: Discrete or continuous; Optimal way to control a dynamic system.
- Prerequisites: Calculus, Vectors and Matrices, ODE and PDE.
- Applications: Production, Finance, Economics, Marketing and others.
Basic Concepts and Definitions

- A dynamic system is described by state equation:

\[ \dot{x}(t) = f(x(t), u(t), t), \quad x(0) = x_0, \tag{1.1} \]

where \( x(t) \) is state variable and \( u(t) \) is control variable.

- The control aim is to maximize the objective function:

\[ J = \int_0^T F(x(t), u(t), t) dt + S[x(T), T]. \tag{1.2} \]

- Usually the control variable \( u(t) \) will be constrained as follows:

\[ u(t) \in \Omega(t), \quad t \in [0, T], \tag{1.3} \]
Sometimes, we consider the following constraints:

1. **Inequality constraints (mixed)**

   \[ g(x(t), u(t), t) \geq 0, \quad t \in [0, T], \quad (1.4) \]

2. **Constraints involving only state variables (pure)**

   \[ h(x(t), t) \geq 0, \quad t \in [0, T], \quad (1.5) \]

3. **Terminal state**

   \[ x(T) \in X, \quad (1.6) \]

   where \( X \) is called the *reachable set* of the state variable at time \( T \).
We now formulate three simple model chosen from areas of production, advertising and economics.

- Production-Inventory Model
- Advertising Model
- Consumption Model
Example 1.1 Production-Inventory Model: The production and inventory storage of a given good to meet an exogenous demand.

<table>
<thead>
<tr>
<th>State Variable</th>
<th>( I(t) = \text{Inventory Level} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variable</td>
<td>( P(t) = \text{Production Rate} )</td>
</tr>
<tr>
<td>State Equation</td>
<td>( \dot{I}(t) = P(t) - S(t), ; I(0) = I_0 )</td>
</tr>
<tr>
<td>Objective Function</td>
<td>Maximize ( J = \int_0^T -[h(I(t)) + c(P(t))] dt )</td>
</tr>
<tr>
<td>State Constraint</td>
<td>( I(t) \geq 0 )</td>
</tr>
<tr>
<td>Control Constraints</td>
<td>( 0 \leq P_{\min} \leq P(t) \leq P_{\max} )</td>
</tr>
<tr>
<td>Terminal Condition</td>
<td>( I(T) \geq I_{\min} )</td>
</tr>
<tr>
<td>Exogenous Functions</td>
<td>( S(t) = \text{Demand Rate} ), ( h(I) = \text{Inventory Holding Cost} ), ( c(P) = \text{Production Cost} )</td>
</tr>
<tr>
<td>Parameters</td>
<td>( T = \text{Terminal Time} ), ( I_{\min} = \text{Minimum Ending Inventory} ), ( P_{\min} = \text{Minimum Possible Production Rate} ), ( P_{\max} = \text{Maximum Possible Production Rate} ), ( I_0 = \text{Initial Inventory Level} )</td>
</tr>
</tbody>
</table>
The Advertising Model

- **Example 1.2 Advertising Model**: A special case of the Nerlove-Arrow advertising model.

<table>
<thead>
<tr>
<th>State Variable</th>
<th>State Equation</th>
<th>Objective Function</th>
<th>State Constraint</th>
<th>Control Constraints</th>
<th>Terminal Condition</th>
<th>Exogenous Function</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(t)$ = Advertising Goodwill</td>
<td>$\dot{G}(t) = u(t) - \delta G(t), G(0) = G_0$</td>
<td>Maximize $\left{ J = \int_0^\infty [\pi(G(t)) - u(t)]e^{-\rho t} dt \right}$</td>
<td>...</td>
<td>$0 \leq u(t) \leq Q$</td>
<td>...</td>
<td>$\pi(G) =$ Gross Profit Rate</td>
<td>$\delta =$ Goodwill Decay Constant</td>
</tr>
<tr>
<td>$u(t)$ = Advertising Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\rho =$ Discount Rate</td>
</tr>
<tr>
<td>$G_0 =$ Initial Goodwill Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$Q =$ Upper Bound on Advertising Rate</td>
</tr>
</tbody>
</table>
### Example 1.3 Consumption Model: A problem of an agent’s consumption of his wealth over time in a way that maximizes his consumption utility over his lifetime.

<table>
<thead>
<tr>
<th>State Variable</th>
<th>$W(t) =$ Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variable</td>
<td>$C(t) =$ Consumption Rate</td>
</tr>
<tr>
<td>State Equation</td>
<td>$\dot{W}(t) = rW(t) - C(t), \enspace W(0) = W_0$</td>
</tr>
<tr>
<td>Objective Function</td>
<td>$\max \left{ \int_0^T U(C(t)) e^{-\rho t} , dt + B(W(T)) e^{-\rho T} \right}$</td>
</tr>
<tr>
<td>State Constraint</td>
<td>$W(t) \geq 0$</td>
</tr>
<tr>
<td>Control Constraint</td>
<td>$C(t) \geq 0$</td>
</tr>
<tr>
<td>Terminal Condition</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>Exogenous Functions</td>
<td>$U(C) =$ Utility of Consumption $\enspace B(W) =$ Bequest Function</td>
</tr>
<tr>
<td>Parameters</td>
<td>$T =$ Terminal Time $\enspace W_0 =$ Initial Wealth $\enspace \rho =$ Discount Rate $\enspace r =$ Interest Rate</td>
</tr>
</tbody>
</table>
History of Optimal Control Theory

- Calculus of variations.
- Brachistochrone problem: path of least time.
- Newton, Leibniz, Bernoulli brothers, Jacobi, Bolza.
Figure 1.1: The Brachistochrone Problem
Notation and Concepts Used

We use the following symbols,

- “=” to mean “is equal to” or “is defined to be equal to” or “is identically equal to”,
- “:=” to mean “is defined to be equal to”,
- “≡” to mean “is identically equal to”,
- “≈” to mean “is approximately equal to”,
- “⇒” to mean “implies”,
- “∀” to mean “for all”,
- “∈” to mean “is a member of”,
- □ indicates the end of a proof.
Let $y$ be an $n$-component column vector and $z$ be an $m$-component row vector, i.e.,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = (y_1, \ldots, y_n)^T \text{ and } z = (z_1, \ldots, z_m),$$

If $y$ and $z$ are functions of time $t$, a scalar, then the time derivatives $\dot{y} := dy/dt$ and $\dot{z} := dz/dt$ are defined as,

$$\dot{y} = \frac{dy}{dt} = (\dot{y}_1, \ldots, \dot{y}_n)^T \text{ and } \dot{z} = \frac{dz}{dt} = (\dot{z}_1, \ldots, \dot{z}_m),$$

When $n = m$, we can define the inner product

$$zy = \sum_{i=1}^{n} z_i y_i. \quad (1.7)$$
More generally, if

\[ A = \{a_{ij}\} = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1k} \\
  a_{21} & a_{22} & \cdots & a_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mk}
\end{bmatrix} \]

is an \( m \times k \) matrix and \( B = \{b_{ij}\} \) is a \( k \times n \) matrix, we define the matrix product \( C = \{c_{ij}\} = AB \), which is an \( m \times n \) matrix with components

\[ c_{ij} = \sum_{r=1}^{k} a_{ir} b_{rj}. \quad (1.8) \]
Differentiating Vectors and Matrices w.r.t. Scalars

Let \( f : E^1 \rightarrow E^k \) be a \( k \)-dimensional function of a scalar variable \( t \). If \( f \) is a row vector, then we define

\[
\frac{df}{dt} = f_t = (f_{1t}, f_{2t}, \cdots, f_{kt}), \text{ a row vector.}
\]

If \( f \) is a column vector, then

\[
\frac{df}{dt} = f_t = \begin{bmatrix} f_{1t} \\ f_{2t} \\ \vdots \\ f_{kt} \end{bmatrix} = (f_{1t}, f_{2t}, \cdots, f_{kt})^T, \text{ a column vector.}
\]
Differentiating Scalars w.r.t. Vectors

If $F(y, z)$ is a scalar function defined on $E^n \times E^m$ with $y$ an $n$-dimensional column vector and $z$ an $m$-dimensional row vector, then the gradients $F_y$ and $F_z$ are defined, respectively, as

$$F_y = (F_{y_1}, \cdots, F_{y_n}), \text{ a row vector,}$$

and

$$F_z = (F_{z_1}, \cdots, F_{z_m}), \text{ a row vector,}$$

where $F_{y_i}$ and $F_{z_j}$ denote the partial derivatives with respect to the subscripted variables.
If $f : E^n \times E^m \to E^k$ is a $k$-dimensional vector function, $f$ either row or column, i.e.,

$$f = (f_1, \cdots, f_k) \text{ or } f = (f_1, \cdots, f_k)^T,$$

where each component $f_i = f_i(y, z)$ depends on the column vector $y \in E^n$ and the row vector $z \in E^m$, then $f_z$ will denote the $k \times m$ matrix

$$f_z = \begin{bmatrix}
\frac{\partial f_1}{\partial z_1}, & \frac{\partial f_1}{\partial z_2}, & \cdots & \frac{\partial f_1}{\partial z_m} \\
\frac{\partial f_2}{\partial z_1}, & \frac{\partial f_2}{\partial z_2}, & \cdots & \frac{\partial f_2}{\partial z_m} \\
\vdots & \vdots & \cdots & \vdots \\
\frac{\partial f_k}{\partial z_1}, & \frac{\partial f_k}{\partial z_2}, & \cdots & \frac{\partial f_k}{\partial z_m}
\end{bmatrix} = \left\{ \frac{\partial f_i}{\partial z_j} \right\},$$

(1.11)
And, $f_y$ will denote the $k \times n$ matrix

$$
f_y = \begin{bmatrix}
\frac{\partial f_1}{\partial y_1}, & \frac{\partial f_1}{\partial y_2}, & \cdots & \frac{\partial f_1}{\partial y_n} \\
\frac{\partial f_2}{\partial y_1}, & \frac{\partial f_2}{\partial y_2}, & \cdots & \frac{\partial f_2}{\partial y_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_k}{\partial y_1}, & \frac{\partial f_k}{\partial y_2}, & \cdots & \frac{\partial f_k}{\partial y_n}
\end{bmatrix} = \left\{ \frac{\partial f_i}{\partial y_j} \right\}. $$

Matrices $f_z$ and $f_y$ are known as *Jacobian* matrices. Thus,

$$f_z = (f^T)_z = f_{zT} = (f^T)_{zT}.$$
Applying the rule (1.11) to $F_y$ in (1.9), we obtain $F_{yz} = (F_y)_z$ to be the $n \times m$ matrix

$$F_{yz} = \begin{bmatrix}
F_{y1z1} & F_{y1z2} & \cdots & F_{y1zm} \\
F_{y2z1} & F_{y2z2} & \cdots & F_{y2zm} \\
\vdots & \vdots & \ddots & \vdots \\
F_{ynz1} & F_{ynz2} & \cdots & F_{ynzm}
\end{bmatrix} = \left\{ \frac{\partial^2 F}{\partial y_i \partial z_j} \right\}. \quad (1.13)$$

Applying the rule (1.12) to $F_z$ in (1.10), we obtain $F_{zy} = (F_z)_y$ to be the $m \times n$ matrix

$$F_{zy} = \begin{bmatrix}
F_{z1y1} & F_{z1y2} & \cdots & F_{z1yn} \\
F_{z2y1} & F_{z2y2} & \cdots & F_{z2yn} \\
\vdots & \vdots & \ddots & \vdots \\
F_{zmy1} & F_{zmy2} & \cdots & F_{zmyn}
\end{bmatrix} = \left\{ \frac{\partial^2 F}{\partial z_i \partial y_j} \right\}. \quad (1.14)$$
Let $g$ be an $n$-component row vector function and $f$ be an $n$-component column vector function of an $n$-component vector $x$. Then in Exercise 1.9, you are asked to show that

$$(gf)_x = gf_x + f^T g_x = gf_x + f^T (g^T)_x.$$  

(1.15)

In Exercise 1.10, you are asked to show further that with $g = F_x$, where $x \in E^n$ and the function $F : E^n \rightarrow E^1$ is twice continuously differentiable so that $F_{xx} = (F_{xx})^T$, called the Hessian, then

$$(gf)_x = (Fxf)_x = Fxf_x + f^T F_{xx} = Fxf_x + (F_{xx} f)^T.$$  

(1.16)
The *norm* of an $m$-component row or column vector $z$ is defined to be

$$\| z \| = \sqrt{z_1^2 + \cdots + z_m^2}.$$  \hfill (1.17)

The *neighborhood* $N_{z_0}$ of a point is defined as,

$$N_{z_0} = \{ z | \| z - z_0 \| < \varepsilon \},$$   \hfill (1.18)

where $\varepsilon > 0$ is a small positive real number.
A function $F(z) : E^m \to E^1$ is said to be of the order $o(z)$, if

$$\lim_{\|z\| \to 0} \frac{F(z)}{\|z\|} = 0.$$ 

The norm of an $m$-dimensional row or column vector function $z(t)$, $t \in [0, T]$, is defined to be

$$\|z\| = \left[ \sum_{j=1}^{m} \int_{0}^{T} z_j^2(\tau) d\tau \right]^{\frac{1}{2}}. \quad (1.19)$$
The Concept of Left & Right Limits

- With $\varepsilon > 0$, these are defined, respectively, for a function $x(t)$ as

  $$x(T^-) = \lim_{\tau \uparrow T} x(\tau) = \lim_{\varepsilon \to 0} x(T - \varepsilon),$$

  $$x(T^+) = \lim_{\tau \downarrow T} x(\tau) = \lim_{\varepsilon \to 0} x(T + \varepsilon).$$

- These limits are illustrated for a function $x(t)$ graphed in Figure 1.2 next.
Here,

\[ x(0) = 1, \]
\[ x(0^+) = 2, \]
\[ x(1^-) = 3, \]
\[ x(1^+) = x(1) = 4, \]
\[ x(2^-) = 3, \]
\[ x(2) = 2, \quad x(2^+) = 1, \]
\[ x(3^-) = 2, \quad x(3) = 3. \]

**Figure 1.2: Illustration of Left and Right Limits**
Some Special Notation

The discrete-time models introduced in Chapter 8 and applied in Chapter 9.

\[ x^k \]  : state variable at time \( k \).

\[ u^k \]  : control variable at time \( k \).

\[ \lambda^k \]  : adjoint variable, respectively at time \( k \).

\[ \Delta x^k \]  := \( x^{k+1} - x^k \) : difference operator.

\( x^{k*} \) and \( u^{k*} \)  : quantities along an optimal path.
Some Special Notation Cont.

- The **bang function**, as

\[
\text{bang}[b_1, b_2; W] = \begin{cases} 
  b_1 & \text{if } W < 0, \\
  \text{arbitrary} & \text{if } W = 0, \\
  b_2 & \text{if } W > 0. 
\end{cases}
\]  

(1.20)

- The **sat function**, as

\[
\text{sat}[y_1, y_2; W] = \begin{cases} 
  y_1 & \text{if } W < y_1, \\
  W & \text{if } y_1 \leq W \leq y_2, \\
  y_2 & \text{if } W > y_2. 
\end{cases}
\]  

(1.21)
Impulse Control, we can compute it’s contribution using

$$\text{imp}(x_1, x_2; t) = \lim_{\varepsilon \to 0} \int_t^{t+\varepsilon} F(x, u, \tau) d\tau.$$  \hfill (1.22)

If the impulse is applied only at time $t$, then we can calculate (1.2) as

$$J = \int_0^t F(x, u, \tau) d\tau + \text{imp}(x_1, x_2; t) + \int_t^T F(x, u, \tau) d\tau + S[x(T), T].$$ \hfill (1.23)
A set $D \subset E^n$ is a convex set if for each pair of points $y, z \in D$, the entire line segment joining these two points is also in $D$, i.e.,

$$py + (1 - p)z \in D, \text{ for each } p \in [0, 1].$$

Given $x^i \in E^n$, $i = 1, 2, \ldots, l$, we define $y \in E^n$ to be a convex combination of $x^i \in E^n$, if there exists $p_i \geq 0$ such that

$$\sum_{i=1}^{l} p_i = 1 \text{ and } y = \sum_{i=1}^{l} p_i x^i.$$

The convex hull of a set $D \subset E^n$ is

$$\text{co}D := \left\{ \sum_{i=1}^{l} p_i x^i : \sum_{i=1}^{l} p_i = 1, p_i \geq 0, x^i \in D, i = 1, 2, \ldots, l \right\}.$$
Concave Function

- \( \psi : D \to E^1 \), is concave, if for each pair of points \( y, z \in D \) and for all \( p \in [0, 1] \),
  \[
  \psi(py + (1 - p)z) \geq p\psi(y) + (1 - p)\psi(z).
  \]

- If \( \geq \) is changed to \( > \) for all \( y, z \in D \) with \( y \neq z \), and \( 0 < p < 1 \), then \( \psi \) is called a strictly concave function.

- If \( \psi(x) \) is a differential function on the interval \([a, b]\), then it is concave, if for each pair of points \( y, z \in [a, b] \),
  \[
  \psi(z) \leq \psi(y) + \psi_x(y)(z - y).
  \]
Concave Function Cont.

Figure 1.3: A Concave Function
Concave and Convex Functions

- If the function $\psi$ is twice differentiable, then it is concave, if at each point in $D$ (i.e., $[a, b]$), the $n \times n$ symmetric matrix $\psi_{xx}$ is negative semidefinite, i.e., all of its eigenvalues.

  \[
  \psi_{xx} \leq 0
  \]

- If $\psi$ defined on a convex set $D \subset E^n$, is concave function, then the negative of the function $\psi$, i.e., $-\psi : D \to E^1$, is a convex function.
Affine Functions and Homogenous Functions of Degree \( k \)

- A function \( \psi : E^n \rightarrow E^1 \) is said to be **affine**, if the function \( \psi(x) - \psi(0) \) is linear.

- Thus, \( \psi \) can be represented as

\[
\psi(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i + b,
\]

where \( a_i, \ i = 1, 2, \ldots, n \), and \( b \) are scalar constants.

- A function \( \psi : E^n \rightarrow E^1 \) is said to be **homogeneous of degree \( k \)**, if \( \psi(bx) = b^k \psi(x) \), where \( b > 0 \) is a scalar constant.

- A linear function \( \psi(x) = ax = \sum_{i=1}^{n} a_i x_i \) is a simple example of a homogeneous function of degree 1.
A point \((\hat{x}, \hat{y}) \in E^n \times E^m\) is called a saddle point of \(\psi(x, y)\), if
\[
\psi(\hat{x}, y) \geq \psi(\hat{x}, \hat{y}) \geq \psi(x, \hat{y}) \quad \text{for all } x \in E^n \text{ and } y \in E^m.
\]

An important concept in two-person zero-sum games is that of a saddle point. Let \(\psi(x, y)\), a real-valued function defined on the space \(E^n \times E^m\), i.e., \(\psi : E^n \times E^m \to E^1\), be the payoff of player 1 and \(-\psi(x, y)\) be the payoff of player 2, when they make decisions \(x\) and \(y\), respectively, in a zero-sum game.
A saddle point may not exist, and even if it exists, it may not be unique. Also,

$$\psi(\hat{x}, \hat{y}) = \max_x \psi(x, \hat{y}) = \min_y \psi(\hat{x}, y).$$

Figure 1.4: An Illustration of a Saddle Point
A set of vectors $a_1, a_2, \ldots, a_m$ in $E^n$, $m \leq n$, is said to be *linearly dependent* if there exist scalars $p_i$ not all zero such that

$$\sum_{i=1}^{m} p_i a_i = 0.$$  \hspace{1cm} (1.24)

If (1.24) holds only when $p_1 = p_2 = \cdots = p_m = 0$, then the vectors are said to be *linearly independent*.

The *rank* of an $m \times n$ matrix $A$, written $\text{rank}(A)$, is the maximum number of linearly independent rows or, equivalently, the maximum number of linearly independent columns of $A$. An $m \times n$ matrix is of *full rank* if

$$\text{rank}(A) = \min\{m, n\}.$$