1 Non-linear Optimization with Constraints:

1.1 Kuhn Tucker Conditions:

The following functions are given:

$$f : \mathbb{R}^n \mapsto \mathbb{R}, \qquad \text{continuously differentiable}$$

$$\vec{g} : \mathbb{R}^n \mapsto \mathbb{R}^m, \qquad \text{continuously differentiable}$$

The **constraint qualification** are fullfilled in point $\vec{z}$, if the gradients $\frac{\partial g_i(\vec{x})}{\partial \vec{x}}|_{\vec{z}}$ of all active constraints (i.e. with $g_i(\vec{z}) = 0$) are linearly independent, or equivalently, if the matrix

$$M = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} & g_1(\vec{z}) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} & 0 & \cdots & g_m(\vec{z}) \end{pmatrix}$$

has full rank (i.e. rank $M = m$).

Now consider the optimization problem

$$\max f(\vec{x}) \text{ s.t. the constraints } \vec{g}(\vec{x}) \geq 0. \tag{1}$$

Kuhn Tucker Theorem:

Let x^* be a solution of the optimization problem (1) with the constraint qualification fullfilled. Then there exists a unique vector of Lagrange multipliers $\vec{\lambda} \in \mathbb{R}^m$ such that the following conditions hold for the Lagrange function

$$L(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \vec{\lambda}^t \vec{g}(\vec{x}) = f(\vec{x}) + \sum_{j=1}^{m} \lambda_j g_j(\vec{x})$$

- $$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\vec{x})}{\partial x_i} = 0, \qquad i = 1, \cdots n \tag{2}$$

- $$\frac{\partial L}{\partial \lambda_j} = g_j(\vec{x}^*) \geq 0 \qquad j = 1, \cdots m \tag{3}$$

- $$\lambda_j \geq 0, \qquad j = 1, \cdots m \tag{4}$$

- $$\vec{\lambda}^t\vec{g}(\vec{x}^*) = \sum_{j=1}^{m} \underbrace{\lambda_j g_j(\vec{x}^*)}_{\geq 0} = 0 \tag{5}$$

As the summands of (5) are all non-negative, condition (5) is equivalent to the m conditions

$$\lambda_j g_j(\vec{x}^*) = 0, \qquad \forall j = 1, \cdots m. \tag{6}$$

Remark: Under the constraint qualification conditions (2)-(4) & (6) are necessary conditions for an optimal solution. If additionally the functions f and g_j are concave, then these conditions are also sufficient.

If the constraint qualification does not hold, then the above conditions are **NOT** necessary.

Interpretation of Lagrange multiplier:

λ_j equals the marginal increase of the objective value, if constraint $g_j(\vec{x}) \geq 0$ is marginally relaxed, i.e. if $\vec{x}^*(\epsilon)$ is the solution of $\max f(\vec{x})$ subject to the constraint $\vec{g}(\vec{x}) \geq -\epsilon$ then

$$\vec{\lambda} = \left.\frac{\partial \vec{f}(\vec{x}^*(\epsilon))}{\partial \epsilon}\right|_{\epsilon=0}$$

2 Optimal Control Theory

A firm wants to maximize its profit over a given planning horizon $[0, T]$. The "state" of the firm is described by a vector of capital values $\vec{x}(t) \in \mathbb{R}^n$. At each instant of time the firm has to decide on its investments, prices, level of production, etc., denoted by $\vec{u}(t) \in \mathbb{R}^m$.

The state $\vec{x}(t)$ together with the control $\vec{u}(t)$ determine the profit at each instant of time and is denoted by $\Pi(\vec{x}(t), \vec{u}(t), t)$.

Integrated over the time horizon $[0, T]$ and discounted at the non-negative discount rate $\rho \geq 0$ the firm wants to maximize

$$J = \int_{t=0}^{T} e^{-\rho t}\Pi(\vec{x}(t), \vec{u}(t), t)dt + e^{-\rho T}S(\vec{x}(T), T) \tag{7}$$

where $S(\vec{x}(T), T)$ denotes the salvage value.

Remark: The objective functional defined by (7) is known as **Bolza-form**. If $S(\vec{x}(T), T) = 0$ the problem is known as **Lagrange problem.** In the case that $\Pi(\vec{x}(t), \vec{u}(t), t) = 0$ the problem is denoted as **Mayer problem.** All 3 formulations are equivalent.

The state evolves according to the differential equation

$$\dot{\vec{x}} = f(\vec{x}(t), \vec{u}(t), t), \quad \text{with given initial state } \vec{x}(0) = \vec{x}_0$$

The problem can be summarized as

$$\max_{\vec{u}(t)\in\Omega\subset\mathbb{R}^m} \{J = \int_{t=0}^{T} e^{-\rho t}\Pi(\vec{x}(t), \vec{u}(t), t)dt + e^{-\rho T}S(\vec{x}(T), T)\}$$

subject to

$$\dot{\vec{x}} = f(\vec{x}(t), \vec{u}(t), t), \quad \vec{x}(0) = \vec{x}_0$$

Assumptions:

- $\Pi(\vec{x}, \vec{u}, t)$ and $f(\vec{x}, \vec{u}, t)$ are continuously differentiable w.r.t. $\vec{x}$ and continuous w.r.t. $\vec{u}$ and t.
- $S(\vec{x}, T)$ is continuously differentiable w.r.t. $\vec{x}$ and T.
- As admissible controls we consider all piecewise continuous functions defined on the time interval $[0, T]$ with values in $\Omega \subset \mathbb{R}^m$.
- Substituting an admissible control trajectory into the state dynamics leads to a continuous, piecewise continuously differentiable state trajectory $\vec{x}(t), t \in [0, T]$.
- If profit Π as well as dynamics f do not explicitly depend on time t, then the system is said to be **autonomous.**

To present Pontryagin's Maximum Principle, we define the current value Hamiltonian

$$\mathcal{H}(\vec{x}, \vec{u}, \vec{\lambda}, t) = \Pi(\vec{x}, \vec{u}, t) + \vec{\lambda}^t f(\vec{x}, \vec{u}, t)$$

where $\vec{\lambda} \in \mathbb{R}^n$ denotes the co-state or adjoint variable.

Maximum Principle: (Theorem 1)

Let $\vec{u}^*(t)$ be the optimal control of the above problem and $\vec{x}^*(t)$ the corresponding state trajectory. The there exists a continuous, piecewise continuously differentiable vector-valued function $\vec{\lambda}(t) \in \mathbb{R}^n$ (denoted as adjoint variables) such that the following statements hold:

- At time points where $\vec{u}^*(t)$ is continuous, the control has to maximize the Hamiltonian, i.e.

$$\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \vec{\lambda}(t), t) = \max_{\vec{u}\in\Omega} \mathcal{H}(\vec{x}^*(t), \vec{u}, \vec{\lambda}(t), t)$$

- The adjoint variable has to follow the differential equation

$$\dot{\vec{\lambda}}(t) = \rho\vec{\lambda}(t) - \frac{\partial\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \vec{\lambda}(t), t)}{\partial\vec{x}}$$

- at the endpoint the transversality condition

$$\vec{\lambda}(T) = \frac{\partial S(\vec{x}^*(T), T)}{\partial\vec{x}}$$

has to hold.

3 "Proof" of *Pontryagin's Maximum Principle* by Dynamic Programming

Define the *Value Function* $V(\vec{x}, t)$ by

$$V(\vec{x}, t) = \max_{\vec{u}(s)\in\Omega} \left\{ \int_t^T e^{-(\rho(s-t)} \Pi(\vec{x}(s), \vec{u}(s), s) ds + e^{-\rho(T-t)} S(\vec{x}(T), T) \right\}$$

Now we consider a short time interval $[t, t+\Delta]$. According to Bellmann's Principle of Optimality we get

$$V(\vec{x}, t) = \max_{\vec{u}(s)\in\Omega} \left\{ \int_t^{t+\Delta} e^{-(\rho(s-t)} \Pi(\vec{x}(s), \vec{u}(s), s) ds + e^{-\rho\Delta} V(\vec{x}(t+\Delta, t+\Delta) \right\}$$

Because of continuity assumptions, maximizing over the time interval $[t, t+\Delta]$ can be approximated by maximizing only at time t, i.e.

$$V(\vec{x}, t) = \max_{\vec{u}(t)\in\Omega} \left\{ \Pi(\vec{x}(t), \vec{u}(t), t)\Delta + e^{-\rho\Delta} V(\vec{x}(t+\Delta, t+\Delta) + o(\Delta) \right\} \tag{8}$$

Due to continuity assumptions, V can be expanded to

$$V(\vec{x}(t+\Delta), t+\Delta) = V(\vec{x}(t), t) + V_{\vec{x}}(\vec{x}(t), t)\dot{\vec{x}}\Delta + V_t(\vec{x}, t)\Delta + o(\Delta) \tag{9}$$

Plugging in (9) into (8) yields

$$V(\vec{x}, t) = \max_{\vec{u}(t)\in\Omega} \{ \Pi(\vec{x}(t), \vec{u}(t), t)\Delta + (1-\rho\Delta)[V(\vec{x}(t), t) + \tag{10}$$

$$+V_{\vec{x}}(\vec{x}(t),t)\dot{\vec{x}}\Delta + V_t(\vec{x},t)\Delta] + o(\Delta)\Big\} \tag{11}$$

where $e^{-\rho\Delta} = 1 - \rho\Delta + o(\Delta)$.

Dividing by Δ and taking $\lim_{\Delta\to 0}$ leads to

$$0 = \max_{\vec{u}(t)\in\Omega} \{\Pi(\vec{x}(t),\vec{u}(t),t) - \rho V(\vec{x}(t),t) + V_{\vec{x}}(\vec{x}(t),t)f(\vec{x},u,t) + V_t(\vec{x},t)\} \tag{12}$$

As boundary condition we have

$$V(\vec{x},T) = S(\vec{x},T) \tag{13}$$

Now define the adjoint variables as

$$\vec{\lambda} = \left(\frac{\partial V(\vec{x}(t),t)}{\partial x_j}\right)_{j=1,\cdots n}$$

Defining the Hamiltonian

$$\mathcal{H}(\vec{x},\vec{u},\vec{\lambda},t) = \Pi(\vec{x},\vec{u},t) + \vec{\lambda}\vec{f}(\vec{x},\vec{u},t)$$

equation (12) can be written as

$$0 = \max_{\vec{u}(t)\in\Omega} \Big\{\mathcal{H}(\vec{x},\vec{u},\vec{\lambda},t) - \rho V(\vec{x}(t),t) + V_t(\vec{x},t)\Big\} \tag{14}$$

which leads to the **Hamilton-Jacobi-Bellman Equation**

$$\rho V(\vec{x}(t),t) + V_t(\vec{x},t) = \max_{\vec{u}(t)\in\Omega} \{\mathcal{H}(\vec{x},\vec{u},V_{\vec{x}},t)\} \tag{15}$$

Obviously, the optimal control has to maximize the Hamiltonian, i.e.

$$\mathcal{H}(\vec{x}^*(t),\vec{u}^*(t),\vec{\lambda}(t),t) \geq \mathcal{H}(\vec{x}^*(t),\vec{u},\vec{\lambda}(t),t)$$

The transversality condition has to hold due to

$$V(\vec{x},T) = S(\vec{x},T) \text{ and } \vec{\lambda} = \left(\frac{\partial V(\vec{x}(t),t)}{\partial x_j}\right)_{j=1,\cdots n}$$

The adjoint equations can be obtained as follows:

Plugging in the optimal solution into equation (14) leads to

$$\mathcal{H}(\vec{x}^*,\vec{u}^*,V_{\vec{x}}(\vec{x}^*,t),t) - \rho V(\vec{x}^*,t) + V_t(\vec{x}^*,t) = 0 \tag{16}$$

which is the maximum value the function

$$\mathcal{H}(\vec{x},\vec{u}^*,V_{\vec{x}}(\vec{x},t),t) - \rho V(\vec{x},t) + V_t(\vec{x},t) \tag{17}$$

can obtain.

For interior solutions we get

$$\mathcal{H}_{\vec{x}}(\vec{x}^*, \vec{u}^*, V_{\vec{x}}(\vec{x}^*, t), t) - \rho V_{\vec{x}}(\vec{x}^*, t) + V_{t\vec{x}}(\vec{x}^*, t) = 0 \tag{18}$$

Differentiation of the Hamiltonian with respect to $\vec{x}$ yields $\mathcal{H}_{\vec{x}} = \Pi_{\vec{x}} + V_{\vec{x}\vec{x}}f + V_{\vec{x}}f_{\vec{x}}$ and therefore

$$\Pi_{\vec{x}} + V_{\vec{x}\vec{x}}f + V_{\vec{x}}f_{\vec{x}} - \rho V_{\vec{x}}(\vec{x}^*, t) + V_{t\vec{x}}(\vec{x}^*, t) = 0 \tag{19}$$

Now

$$\frac{dV_{\vec{x}}}{dt} = V_{\vec{x}\vec{x}}\dot{\vec{x}} + V_{\vec{x}t} = \rho V_{\vec{x}} - \Pi_{\vec{x}} - V_{\vec{x}}f_{\vec{x}}$$

Now assuming that $\vec{\lambda}$ only depends on time and not on the state, we obtain the adjoint equation

$$\dot{\vec{\lambda}} = \rho\vec{\lambda} - \Pi_{\vec{x}} - \vec{\lambda}f_{\vec{x}}$$

Extensions

Consider the optimal control problem

$$\max_{\vec{u}(t)\in\Omega\subset\mathbb{R}^m} \{J = \int_{t=0}^{T} e^{-\rho t}\Pi(\vec{x}(t), \vec{u}(t), t)dt + e^{-\rho T}S(\vec{x}(T), T)\}$$

subject to

$$\dot{\vec{x}} = f(\vec{x}(t), \vec{u}(t), t), \quad \vec{x}(0) = \vec{x}_0$$

with additional terminal conditions $(0 \leq n_1 \leq n_2 \leq n)$:

$$\begin{array}{lcll} x_j(T) & \cdots & \text{arbitrary} \ , & j = 1, \cdots, n_1 \\ x_j(T) & = & x_j^T & j = n_1 + 1, \cdots, n_2 \\ x_j(T) & \geq & x_j^T & j = n_2 + 1, \cdots, n \end{array}$$

We define the Hamiltonian as

$$\mathcal{H}(\vec{x}, \vec{u}, \lambda_0, \vec{\lambda}, t) = \lambda_0\Pi(\vec{x}, \vec{u}, t) + \vec{\lambda}^t f(\vec{x}, \vec{u}, t)$$

where λ_0 is a non-negative constant.

Remark: For the standard problem (normal case) the constant can be set to $\lambda_0 = 1$. Under the assumption of additional terminal conditions, the "abnormal" case $\lambda_0 = 0$ cannot be excluded a priori.

Maximum Principle for the standard problem with terminal conditions:

Let $\vec{u}^*(t)$ be the optimal control of the above problem and $\vec{x}^*(t)$ the corresponding state trajectory. The there exists a constant $\lambda_0 \geq 0$ and a continuous, piecewise continuously differentiable vector-valued function $\vec{\lambda}(t) \in \mathbb{R}^n$ (denoted as adjoint variables) such that the following statements hold:

- $(\lambda_0, \vec{\lambda}(t)) \neq (0, \vec{0}) \forall t \in [0, T]$
- At time points where $\vec{u}^*(t)$ is continuous, the control has to maximize the Hamiltonian, i.e.

$$\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \lambda_0, \vec{\lambda}(t), t) = \max_{\vec{u} \in \Omega} \mathcal{H}(\vec{x}^*(t), \vec{u}, \lambda_0, \vec{\lambda}(t), t)$$

- The adjoint variable has to follow the differential equation

$$\dot{\vec{\lambda}}(t) = \rho \vec{\lambda}(t) - \frac{\partial \mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \lambda_0, \vec{\lambda}(t), t)}{\partial \vec{x}}$$

- at the endpoint the following transversality conditions have to hold

$$\begin{array}{lll} \lambda_j(T) = \lambda_0 \frac{\partial S(\vec{x}^*(T),T)}{\partial x_j} & \text{for} & j = 1, \cdots, n_1 \\ \lambda_j(T) \text{ arbitrary} & \text{for} & j = n_1 + 1, \cdots, n_2 \\ \lambda_j(T) \geq \lambda_0 \frac{\partial S(\vec{x}^*(T),T)}{\partial x_j} & \text{for} & j = n_2 + 1, \cdots, n \\ [\lambda_j(T) - \lambda_0 \frac{\partial S(\vec{x}^*(T),T)}{\partial x_j}][x_j^*(T) - x_j^T] = 0 & & \end{array}$$

Sufficiency Conditions

Up to now we only dealt with candidates for optimal solutions, as the above theorems only give necessary conditions. There are 3 possibilities to make sure that a candidate is indeed optimal:

1. one proves the existence of an optimal solution and that the candidate is the only one which fullfils the necesarry optimality conditions.

2. For a given feasible solution find the value function $V(\vec{x}, t)$ and show that the Hamilton-Jacobi-Bellman equation holds for this solution.

3. Show that additional concavity assumptions hold for the Hamiltonian.

Necessary optimality conditions for the standard problem:

Let $\vec{u}^*(t)$ be a feasible control of the standard problem and $\vec{x}^*(t)$ the corresponding state trajectory. Moreover there exists a trajectory of adjoint variables $\vec{\lambda}(t) \in \mathbb{R}^n$ such that the conditions of the maximum principle hold, i.e.

- $$\dot{\vec{x}}^* = f(\vec{x}^*, \vec{u}^*, t), \qquad \vec{x}(0) = \vec{x}_0$$

- $$\dot{\vec{\lambda}}(t) = \rho\vec{\lambda}(t) - \frac{\partial\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \vec{\lambda}(t), t)}{\partial\vec{x}}, \qquad \vec{\lambda}(T) = \frac{\partial S(\vec{x}^*(T), T)}{\partial\vec{x}}$$

- $$\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \vec{\lambda}(t), t) = \max_{\vec{u}\in\Omega}\mathcal{H}(\vec{x}^*(t), \vec{u}, \vec{\lambda}(t), t) = \mathcal{H}^\circ(\vec{x}^*(t), \vec{\lambda}(t), t)$$

Define the maximized Hamiltonian as

$$\mathcal{H}^\circ(\vec{x}, \vec{\lambda}(t), t) = \max_{\vec{u}\in\Omega}\mathcal{H}(\vec{x}, \vec{u}, \vec{\lambda}(t), t)$$

If the maximized Hamiltonian $\mathcal{H}^\circ(\vec{x}, \vec{\lambda}(t), t)$ is concave and continuously differentiable w.r.t. $\vec{x}$ for all $(\lambda(t), t)$ and if the salvage value $S(\vec{x}, T)$ is concave w.r.t. $\vec{x}$, then the control $\vec{u}^*$ is optimal, i.e. the conditions of the maximum principle are also sufficient.

In case that $\mathcal{H}^\circ$ is stricktly concave, the optimal solution is unique.

Infinite time horizon

We now consider problems of the form

$$\max_{\vec{u}(t)\in\Omega\subset\mathbb{R}^m}\{J = \int_{t=0}^{\infty} e^{-\rho t}\Pi(\vec{x}(t), \vec{u}(t), t)dt\}$$

subject to

$$\dot{\vec{x}} = f(\vec{x}(t), \vec{u}(t), t), \quad \vec{x}(0) = \vec{x}_0$$

Assumptions:

- $\Pi(\vec{x}, \vec{u}, t)$ and $f(\vec{x}, \vec{u}, t)$ are continuously differentiable w.r.t. $\vec{x}$ and continuous w.r.t. $\vec{u}$ and t.

- $S(\vec{x}, T)$ is continuously differentiable w.r.t. $\vec{x}$ and T.

- As admissible controls we consider all piecewise continuous functions defined on the time interval $[0, T]$ with values in $\Omega \subset \mathbb{R}^m$.
- Substituting an admissible control trajectory into the state dynamics leads to a continuous, piecewise continuously differentiable state trajectory $\vec{x}(t), t \in [0, T]$.
- The integral converges for all admissible solutions.

We define the Hamiltonian as

$$\mathcal{H}(\vec{x}, \vec{u}, \lambda_0, \vec{\lambda}, t) = \lambda_0 \Pi(\vec{x}, \vec{u}, t) + \vec{\lambda}^t f(\vec{x}, \vec{u}, t)$$

Theorem: Maximum Principle for infinite time horizon

Let $(\vec{u}^*(t), \vec{x}^*(t))$ be a pair of feasible control and state trajectories. Necessary for the optimality of the pair $(\vec{u}^*(t), \vec{x}^*(t))$ is the existence of a constant $\lambda_0 \geq 0$ and of a continuous co-state trajectory $\vec{\lambda}(t)$ such that $(\lambda_0, \vec{\lambda}(t)) \neq (0, \vec{0}) \forall t \in [0, \infty)$ and that the following conditions hold:

- At time points where $\vec{u}^*(t)$ is continuous, the control has to maximize the Hamiltonian, i.e.
$$\mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \lambda_0, \vec{\lambda}(t), t) = \max_{\vec{u} \in \Omega} \mathcal{H}(\vec{x}^*(t), \vec{u}, \lambda_0, \vec{\lambda}(t), t)$$
- The adjoint variable has to follow the differential equation
$$\dot{\vec{\lambda}}(t) = \rho \vec{\lambda}(t) - \frac{\partial \mathcal{H}(\vec{x}^*(t), \vec{u}^*(t), \lambda_0, \vec{\lambda}(t), t)}{\partial \vec{x}}$$

Remark:

- The corresponding transversality condition
$$\lim_{t \to \infty} e^{-\rho t} \vec{\lambda}(t) = \vec{0}$$
is **NOT** a necessary condition for optimality.

example:

Maximize $J = \int_0^\infty (1-x) u dt$, subject to $\dot{x} = (1-x)u$, $x(0) = 0, 0 \leq u \leq 1$.

Obviously we have

$$J = \int_0^\infty (1-x) u dt = \int_0^\infty \dot{x} dt = \lim_{t \to \infty} x(t) - x(0) = x(\infty).$$

Each feasible solution which maximizes $x(\infty)$ is optimal. The maximum value for $x(\infty)$ is $x(\infty) = 1$.

For example $u^* = 0.5$ is an optimal solution.

The Hamiltonian is $\mathcal{H} = (\lambda_0 + \lambda)(1 - x)u$

and therefore

$$u^* = \begin{cases} 0 & \text{for} \quad \mathcal{H}_u < 0 \\ undefined & \text{for} \quad \mathcal{H}_u = 0 \\ 1 & \text{for} \quad \mathcal{H}_u > 0 \end{cases}$$

For the optimal solution $u^* = 0.5$ therefore we have $\mathcal{H}_u = (\lambda_0 + \lambda)(1 - x) = 0$ and thus $\lambda_0 = -\lambda$. $\lambda < 0$ because of $(\lambda_0, \lambda) \neq (0, 0)$ and $\lambda_0 \geq 0$ violating the transversality condition $\lim_{t\to\infty} \lambda = 0$ (Note that $\rho = 0$.

- Contrary to the standard problem with finite time horizon the constant λ_0 cannot be set to $\lambda_0 = 1$ *a priori*.

  Example:

  Maximize $J = \int_0^\infty (u - x)dt$, subject to $\dot{x} = u^2 + x$, $x(0) = 0, 0 \leq u \leq 1$.

  If $u \neq 0$ on a time interval of non-zero length, then x will diverge to ∞ and $J = -\infty$ as $u \leq 1$. Therefore the optimal solution is $u^* = 0$.

  The Hamiltonian is $\mathcal{H} = \lambda_0(u - x) + \lambda(u^2 + x)$

  Consider the derivative of the Hamiltonian w.r.t. u at $u = 0$.

$$\left.\frac{\partial \mathcal{H}}{\partial u}\right|_{u=0} = \lambda_0 + 2\lambda u|_{u=0} = \lambda_0 \geq 0$$

  In case that $\lambda_0 \neq 0$ the Hamiltonian is increasing at $u = 0$ and therefore u cannot maximize the Hamiltonian, which leads to a contradiction. Therefore $\lambda_0 = 0$.

3.1 Optimal choice of terminal time T

ex: optimal maintanance and optimal choice of time to sell the machine.

Consider the optimal control problem with finite terminal time where additionally the terminal time has to be chosen optimally to maximize

$$\max_{\vec{u}(t)\in\Omega\subset\mathbb{R}^m}\{J=\int_{t=0}^{T}e^{-\rho t}\Pi(\vec{x}(t),\vec{u}(t),t)dt+e^{-\rho T}S(\vec{x}(T),T)\}$$

Additionally to the necessary conditions of the Maximum Principle (Theorem 1) we have

$$\frac{dJ}{dT}|_{T=T^*}=0$$

Taking the derivative of the objective functional w.r.t. T leads to

$$e^{-\rho T}\Pi+e^{-\rho T}(S_{\vec{x}}\dot{x}+S_T-\rho S)=0$$

This implies

$$\mathcal{H}(\vec{x}^*(T^*),\vec{u}^*(T^*),\vec{\lambda}(T^*),T^*)=\Pi+\vec{\lambda}f=$$

$$=\Pi+S_{\vec{x}}\dot{\vec{x}}=\rho S(\vec{x}^*(T^*),T^*)+S_T(x^*(T^*),T^*)$$

linear optimal control models

We now consider optimal control models with the property that the Hamiltonian is linear in the control u. To avoid unbounded values for the control we have to assume, that $u(t)\in[\underline{u},\bar{u}]$.

Define the **switching function** $\sigma(t)=\mathcal{H}_u$. The optimal control is then given as

$$u(t)=\begin{cases}\underline{u} & \text{if} \quad \sigma(t)<0\\ \text{undefined} & \text{if} \quad \sigma(t)=0\\ \bar{u} & \text{if} \quad \sigma(t)>0\end{cases}$$

A solution is called a **bang-bang solution** if the optimal control only takes the values $\underline{u}$ and $\bar{u}$.

If $\sigma(t)=0$ for all $t\in[t_1,t_2]$ $(t_1<t_2)$, then the optimal control can take values in the interval $[\underline{u},\bar{u}]$. Such parts of the optimal solution are called **singular arcs.**

Along the singular arc we have $\sigma(t)=\dot{\sigma}(t)=0$ for all $t\in[t_1,t_2]$. From this condition the optimal control can be computed.