6

OPTIMAL CONTROL in CONTINUOUS TIME

- CONTINUOUS OPTIMAL CONTROL
  - Formulation(s)
  - Multistage problems
  - (Hybrid problems)
  - Numerical technique
- NUMERICAL INTEGRATION OF DYNAMIC SYSTEMS
  - Explicit one-step
  - Stiff systems and implicit integration
  - Orthogonal collocation
  - (1^{st} and 2^{nd} Order Sensitivities)
- HAMILTON-JACOBI-BELLMAN
  - Dynamic Programming in continuous time
  - Linear Quadratic Control and Riccati Equ.
  - Infinite-time optimal control
- PONTRYAGIN and the indirect approach
  - HJB along the optimal solution
  - Controls on regular and singular arcs
  - Pontryagin with path constraints
  - (Hamiltonian systems)
  - (Calculus of variations, connection)
  - Numerical Solution
- DIRECT APPROACHES
  - Direct Single-, Direct-multiple shooting
  - Direct Collocation

CONTINUOUS TIME OPTIMAL CONTROL

6a

The general setup refers to the case in which both the state variables and the input variables are continuous

→ FUNCTIONS, THUS ELEMENTS OF A ∞-DIMENSIONAL FUNCTIONAL SPACE

WE DISCUSS HOW TO FORMULATE THIS CONTINUOUS-TIME OPTIMAL CONTROL PROBLEM AND HOW TO CONVERT IT INTO A DISCRETE-TIME OPTIMAL-CONTROL PROBLEM THAT CAN BE SOLVED WITH A COMPUTER

WE CAN CONSIDER THE FOLLOWING FORMULATION

$$\underset{x(\cdot),\,u(\cdot)}{\text{minimize}} \int_0^T \mathcal{L}(x(t), u(t))\,dt + E(x(T))$$

USED TO BE VECTORS, THEY ARE NOW FUNCTIONS (infinite-dimensional vectors)

USED TO BE A SUM, IT IS NOW AN INTEGRAL (infinite number of term)

USED TO BE THE INDEX k (infinite number of time steps)

subject to

- $x(0) - x_0 = 0$
- $\dot{x}(t) - f(x(t), u(t)) = 0$ for $t \in [0, T]$
- $h(x(t), u(t)) \leq 0$ for $t \in [0, T]$

a constant zero in time (a function)

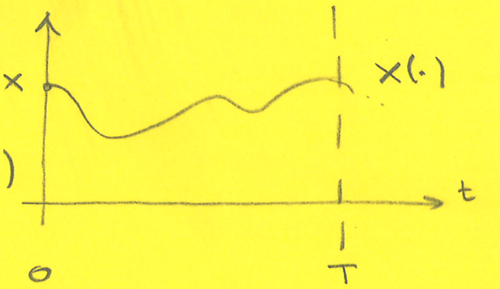
- $r(x(T)) \leq 0$

TERMINAL CONSTRAINTS

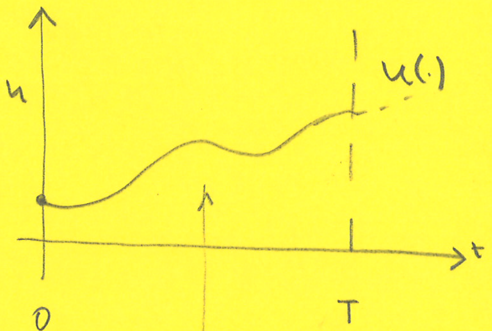
WE HAVE OMITTED OTHER EQUALITY CONSTRAINTS

WE OMITTED TIME DEPENDENCE FOR SIMPLICITY

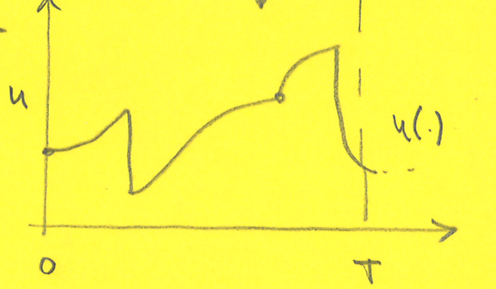
$x(\cdot)$ STATES ARE FUNCTION OF TIME IN SOME INTERVAL



$u(\cdot)$ INPUTS ARE ALSO FUNCTION OF TIME



(this may also have jumps)



CTOC 1

$$\underset{x(\cdot),u(\cdot)}{\text{minimise}} \quad \underbrace{\underbrace{\int_0^T \mathcal{L}(x(t),u(t))\,dt}_{\text{LAGRANGE TERM}} + \underbrace{E(x(T))}_{\text{MAYER TERM}}}_{\text{ALTOGETHER, BOLZA OBJECTIVE}}$$

s.t

$$\dot{x}(t) - f(x(t),u(t)) = 0 \qquad t \in [0,T]$$

$$h(x(t),u(t)) \leq 0 \qquad t \in [0,T]$$

In principle, all functions ($\mathcal{L}$, h, r and E) can be function of time t

→ WE DO NOT CONSIDER THIS GENERALISATION

→ IT IS HANDLED BY INTRODUCING A 'CLOCK STATE' τ

The augmented state $\tilde{x} = \begin{bmatrix} x(t) \\ \tau(t) \end{bmatrix}$ such that $\dot{\tilde{x}} = \begin{bmatrix} f(x,u,t) \\ 1 \end{bmatrix}$

$\tilde{x}(0) = \begin{bmatrix} x(0) \\ 0 \end{bmatrix}$

The differential equation for passing of time and the beginning of time

In principle, also the duration T can be treated as decision variable

→ FREE INITIAL TIME

→ FREE FINAL TIME

We can rescale the time horizon to the interval [0,1] by a time constant but free variable T, to be optimised

The augmented state

$$\tilde{x} = \begin{bmatrix} x \\ T \end{bmatrix}, \qquad \tilde{f}(x,u) = \begin{bmatrix} T f(x,u) \\ 0 \end{bmatrix}, \qquad \text{with pseudo time } \tau \in [0,1]$$

T IS TREATED AS A PARAMETER, THE INITIAL CONDITION T(0) FOR THE STATE T IS FREE AND T SATISFIES $\dot{T} = 0$

MULTISTAGE PROBLEMS

THESE PROBLEMS DEFINE A SPECIAL CLASS OF OPTIMAL CONTROL PROBLEMS IN WHICH THE PROBLEM FORMULATION CHANGES DURING THE INTERVAL

→ DYNAMICS, COST FUNCTION AND CONSTRAINTS CHANGE DISCONTINUOUSLY AT SOME TIME POINT

→ THE TIME OF THE SWITCH CAN BE FIXED, FREE OR EVENT RELATED

→ WE ASSUME THAT THE ORDERING OF THE SWITCHES IS FIXED

WE CONSIDER EACH STAGE AS AN OPTIMAL CONTROL PROBLEM AND WE LINK STAGES BY MATCHING THE FINAL STATE OF ONE STAGE TO THE INITIAL CONDITION OF THE NEXT STAGE

$$\underset{x(\cdot),\,u(\cdot),\,T_1,T_2,\ldots,T_N}{\text{minimize}} \quad \sum_{k=0}^{N-1}\left[\int_{T_k}^{T_{k+1}} \mathcal{L}_k(x_k(t),u_k(t))\,dt + E_k(x_k(T_{k+1}))\right]$$

subject to

$$x_0(T_0) - \bar{x}_0 = 0 \qquad \text{(initial value)}$$

$$x_k(T_k) - x_{k-1} = 0 \qquad \text{(State continuity)}$$

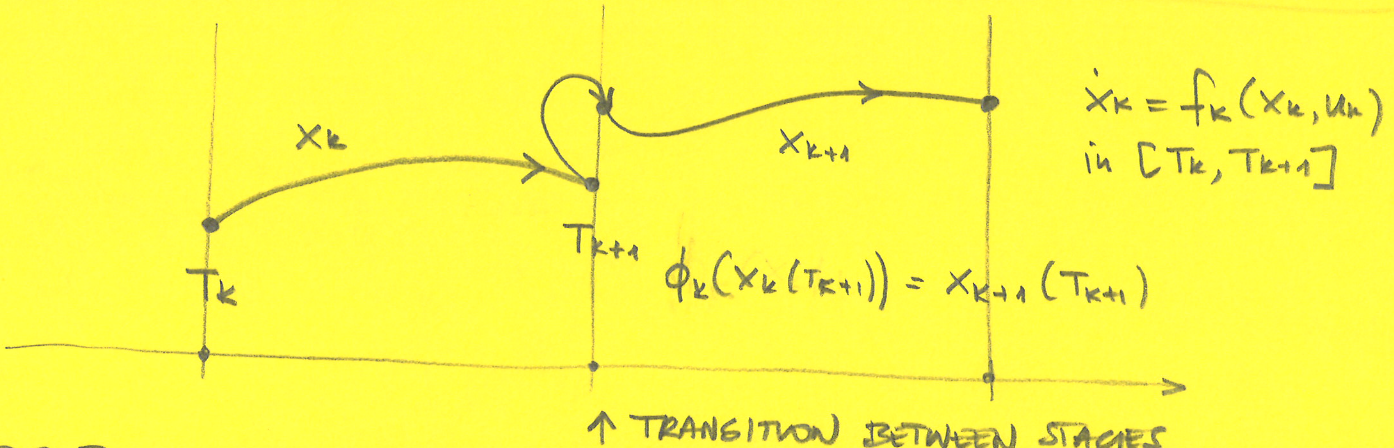
$$\dot{x}_k(t) - f_k(x_k(t),u_k(t)) = 0 \qquad \text{(Dynamics)} \qquad t\in[T_k,T_{k+1}]$$

$$h_k(x_k(t),u_k(t)) \le 0 \qquad \text{(Path constraints)} \qquad t\in[T_k,T_{k+1}]$$

$$r_k(x_k(T_k)) \le 0 \qquad \text{(terminal constraint)}$$

$$T_k - T_{k+1} \le 0 \qquad \text{(Ordering in time)}$$

N-STAGE PROBLEM



CONTINUOUS-TIME OPTIMAL CONTROL

HAMILTON-JACOBI BELLMAN

(Tabulation in SS)

Relates to Dynamic Programming in continuous time

INDIRECT METHODS

(Pontryagin, BVP)

Typical for continuous time (First optimize then discretize)

DIRECT METHODS

(NLP solut.)

First discretize then optimize

* DIRECT SINGLE SHOOTING

- Discretized controls in NLP (sequential)
- Simulated dynamics

TRANSCRIPTION

* DIRECT COLLOCATION

- Discretized controls and states in NLP (Simultaneous)

* DIRECT MULTIPLE SHOOTING

- Control and node start values in NLP (Simultaneous)
- Simulate the dynamics over short intervals (may lead to discontinuities then enforce continuity)

NUMERICAL APPROACHES

THREE BIG FAMILIES OF APPROACHES TO ADDRESS THE CTOC TASK

→ STATE SPACE (HJB)

→ INDIRECT (PONTRYAGIN)

→ DIRECT (NLP)

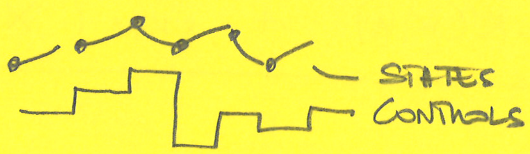
* STATE-SPACE METHODS (HJB) ⇝ RELATED TO DYNAMIC PROGRAMMING
* INDIRECT METHODS (PONTRYAGIN) ⇝ VARIATIONAL CALCULUS
* DIRECT METHODS (NLP) ⇝ BASED ON DISCRETIZATION

'OPTIMIZE, THEN DISCRETIZE'

'DISCRETIZE, THEN OPTIMIZE'

SINGLE SHOOTING : Discretize the controls, then simulates the system dynamics (thus, sequential approach)

MULTIPLE SHOOTING : Discretize the controls and the state variables
(simultaneous)



(the states are simulated over short intervals, independently)

Possibile discontinuities then enforces continuity

DIRECT TRANSCRIPTION METHODS :
(COLLOCATION)
(simultaneous)

Discretize everything

EXPLICIT EULER'S METHOD, as a possibility

$$x_{k+1} = x_k + h\, f(x_k, u_k)$$

IN DYNAMIC PROGRAMMING WE HAVE THE COST-TO-GO AT EACH TIME STEP

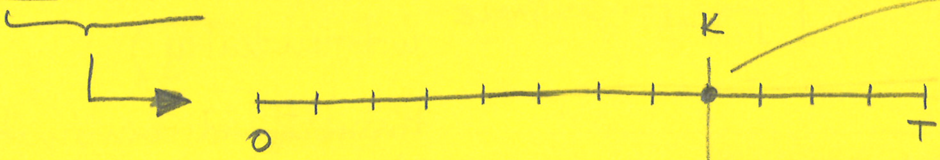
$J_0, J_1, \_\_\_\_, J_{N-1}, J_N$

WITH $\boxed{J_N(x_N) = E(x_N)}$

THEN WE HAVE THE DYNAMIC PROGRAMMING RECURSION, $k = N-1, \_\_\_, 0$

$$\boxed{J_k(x) = \min_u \; h\,\ell(x,u) + J_{k+1}\left(x + h\, f(x,u)\right)}$$

CORRESPONDS TO t_k



$J_k(x) \triangleq J(x, t_k)$

WE CAN START FROM A DISCRETIZATION ACCORDING TO EULER OF THE OPTIMAL CONTROL PROBLEM

$$\boxed{\begin{array}{lll} \underset{x(\cdot),\,u(\cdot)}{\text{minimise}} & \int_0^T \mathcal{L}(x(t),u(t)) + E(T) & \\ \text{subject to} & x(0) - \bar{x}_0 = 0 & \text{(Fixed initial state)} \\ & \dot{x}(t) - f(x(t),u(t)) = 0 \quad t \in [0,T] & \text{(Dynamics)} \end{array}}$$

The choice of Euler's discretization is only illustrative because hardly used in practice

WE START BY INTRODUCING A TIMESTEP $h = \frac{T}{N}$ AND CONSIDER

$$\boxed{\begin{array}{ll} \underset{x,\,u}{\text{minimise}} & \sum_{i=0}^{N-1} h\mathcal{L}(x_i,u_i) + E(x_N) \\ \text{subject to} & x_0 - \bar{x}_0 = 0 \\ & x_{i+1} - x_i - hf(x_i,u_i) = 0 \quad i = 0,1,\dots,N-1 \end{array}}$$

DYNAMIC PROGRAMMING IS APPLIED TO THIS OPTIMISATION PROBLEM

$\rightsquigarrow$ WE HAVE $\boxed{J_K(x) = \underset{u}{\text{minimise}}\; h\mathcal{L}(x,u) + J_{K+1}(x + hf(x,u))}$

THE DYNAMIC PROGRAMMING RECURSION

WE THEN REPLACE THE INDEX K BY ACTUAL TIME POINTS $t_K = Kh$

$$J(x,t_K) = \underset{u}{\text{minimise}}\; h\mathcal{L}(x,u) + J(x + hf(x,u), t_K + h)$$

ASSUMING DIFFERENTIABILITY OF $J(x,t)$ WRT x,t, WE CAN EXPAND IT USING ITS TAYLOR SERIES TRUNCATED TO FIRST ORDER

$$J(x,t) = \text{minimise}\; h\mathcal{L}(x,u) + J(x,t) + h\nabla_x J(x,t)^T f(x,u) + h\frac{\partial}{\partial t}J(x,t) + O(h^2)$$

ABOUT THE TAYLOR'S EXPANSION WRT x AND t OF THE $J(x,t)$

* $J_k(x) = \underset{u}{\text{minimise}}\ h\mathcal{L}(x,u) + J_{k+1}(x + h f(x,u))$

* WE LET $t_k = kh \rightsquigarrow J_k(x) = J(x, t_k)$

* SUBSTITUTING $\rightsquigarrow J(x,t_k) = \underset{u}{\text{minimise}}\ h\mathcal{L}(x,u) + \underbrace{J(x+hf(x,u), t_k+h)}$

Assuming differentiability of $J(x,t)$ wrt (x,t), we write:

$$J(x,t_k) = \underset{u}{\text{minimise}}\ h\mathcal{L}(x,u) + J(x,t_k) + h\nabla_x J(x,t_k)^T f(x,u) + h\frac{\partial J}{\partial t}(x,t_k) + \text{H.O.T.}$$

WE CAN MOVE ALL THE TERMS THAT ARE INDEPENDENT OF u TO THE LEFT

$$\cancel{J(x,t)} - h\frac{\partial J}{\partial t}(x,t) = \underset{u}{\text{minimise}}\ h\mathcal{L}(x,u) + h\nabla_x J(x,t)^T f(x,u) + \cancel{J(x,t)}$$

$$\boxed{\rightsquigarrow -\frac{\partial J}{\partial t}(x,t) = \text{minimise}\ \mathcal{L}(x,t) + \nabla_x J(x,t)^T f(x,u)}$$

HJB

Hamiltonian function

$H(x, \nabla_x J, u)$

$= L(x,u) + \underline{\lambda^T} f(x,u)$

$(t_{k+1} = t_k + h)$

BY BRINGING ALL TERM INDEPENDENT OF u TO THE LEFT HAND SIDE AND BY DIVIDING BY $h \to 0$, WE OBTAIN THE <u>HAMILTON-JACOBI-BELLMAN (HJB) EQUATION</u>

$$\boxed{-\frac{\partial}{\partial t} J(x,t) = \underset{u}{\text{minimise}} \; \mathcal{L}(x,u) + \nabla_x J(x,t)^T f(x,u)}$$

THIS PARTIAL DIFFERENTIAL EQUATION DESCRIBES THE EVOLUTION IN TIME OF THE VALUE FUNCTION

⇝ We solve it backwards for $t \in [0,T]$ starting at the end of the horizon with $J(x,t) = E(x)$

THE OPTIMAL FEEDBACK CONTROL FOR THE STATE x AT TIME t IS THEN OBTAINED AS

$$\boxed{u^*(x,t) = \underset{u}{\arg\min} \; \mathcal{L}(x,u) + \nabla_x J(x,t)^T f(x,u)}$$

IT DEPENDS ON THE GRADIENT OF THE VALUE FUNCTION

$$\nabla_x J(x,t)^T$$

- WE CAN THINK OF IT AS <u>HAMILTONIAN FUNCTION</u>

$$H(x,u,\lambda) \triangleq \mathcal{L}(x,u) + \lambda^T f(x,u)$$

- $u^*(x,t) = \underset{u}{\arg\min} \; H(x,u,\lambda)$

THE 'TRUE HAMILTONIAN' IS THEN COMPUTED

$$H^*(x,\lambda) \triangleq \min_u H(x,u,\lambda) = H(x,\lambda,u^*(x,\lambda))$$

NOTE THAT THE CONTROL DOES NOT APPEAR EXPLICITLY ANYLONGER

LIKE DISCRETE-TIME DISCRETE-STATE DYNAMIC PROGRAMMING, ALSO THE SOLUTION OF THE HJB EQUATION SUFFERS FROM THE CURSE OF DIMENSIONALITY

→ IT'S NUMERICAL SOLUTION IS EXPENSIVE IN MODERATELY LARGE STATE DIMENSIONS, BECAUSE OF THE NEED TO SOLVE A LARGE SIZE DIFFERENTIAL EQUATION

→ ALSO, THE DIFFERENTIABILITY OF THE VALUE FUNCTION IS NOT GUARANTEED ALWAYS

→ IT MAKES IT DIFFICULT TO PROVE EVEN THE EXISTANCE OF A SOLUTION

THE LINEAR-QUADRATIC CASE CAN BE SOLVED

LINEAR - QUADRATIC CONTROL (+ RICCATI EQUATIONS)

The general formulation of the problem

$$\begin{aligned}
&\underset{x(\cdot),\, u(\cdot)}{\text{minimise}} && \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q(t) & S(t)^T \\ S(t) & R(t) \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + x(t)^T P_T x(t) \\
&\text{subject to} && x(0) - x_0 = 0 \qquad \text{(FIXED INITIAL STATE)} \\
& && \dot{x}(t) - A_t x(t) - B_t u(t) = 0 \qquad \text{(LTV DYNAMICS)}
\end{aligned}$$

As in the discrete-time case, the value function is quadratic

⇝ IT CAN BE VERIFIED BY OBSERVING THAT $J(x,T) = x^T P_T x$ IS QUADRATIC

⇝ WE CAN ASSUME (FOR NOW) THAT THIS HOLDS TRUE OVER THE ENTIRE TIME INTERVAL ($J(x,t)$ IS QUADRATIC FOR $t \in [0,T]$ AND HAS THE FORM $x^T P(t) x$ FOR SOME MATRIX $P(t)$)

Under this assumption, the HJB equation becomes

$$-\partial/\partial t \, J(x,t) = \underset{u}{\text{minimise}} \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q(t) & S(t)^T \\ S(t) & R(t) \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + $$

$$+ 2x^T P(t)(A(t)x + B(t)u)$$

If symmetrised, the right hand side reads

$$\underset{u}{\text{minimise}} \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q + PA + A^T P & S^T + PB \\ S + B^T P & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

By the Schur's complement lemma, we get

$$-\partial/\partial t \, J(x,t) = x^T \left(Q + PA + A^T P - (S + PB) R^{-1} (S + B^T P) \right) x$$

which is again a quadratic term.

$$-\partial/\partial t \; J(x,t) = x^T \left(Q + PA + A^T P - (S^T + PB) R^{-1} (S + B^T P) \right) x$$

So $J(x,T)$ is quadratic and remains so for all $t \in [0,T]$ throughout the backwards evolution

$\longrightarrow$ THE RESULTING MATRIX DIFFERENTIAL EQUATION IS

$$\boxed{\begin{array}{l} -\dot{P} = Q + PA + A^T P - (S^T + PB) R^{-1} (S + B^T P)\, x \\ \\ \text{WITH TERMINAL CONDITION } P(T) = P_T \end{array}}$$

DIFFERENTIAL
RICCATI EQUATION

Integrating it backwards allows us to compute the cost-to-go function

The corresponding feedback law is by the Schur's complement lemma given by

$$u^*(x,t) = \underbrace{-R^{-1}(t)\left[S(t) + B^T(t)\, P(t)\right]}_{\text{THE LQR FEEDBACK GAIN}} x$$

INFINITE HORIZON OPTIMAL CONTROL

IF WE CONSIDER THE SAME OPTIMAL CONTROL PROBLEM (LTV dynamics and quadratic cost) OVER AN INFINITE-TIME HORIZON WE HAVE

$$\boxed{\begin{aligned} &\underset{x(\cdot),u(\cdot)}{\text{minimise}} \quad \int_0^\infty \mathcal{L}(x(t),u(t))\,dt \\ &\text{subject to} \quad x(0) - x_0 = 0 \\ &\qquad\qquad\quad \dot{x}(t) - f(x(t),u(t)) = 0 \,, \quad t \in [0,\infty] \end{aligned}}$$

By the principle of optimality, the value function of this problem (if it exists and it is finite) must be stationary

thus by setting $-\partial J/\partial t = 0$, we get

$$\underbrace{0 = \underset{u}{\text{minimise}} \; \mathcal{L}(x,u) + \nabla_x J(x)^T f(x,u)}_{\text{THE STATIONARY HJB EQUATION}}$$

with STATIONARY OPTIMAL CONTROL LAW

$$\boxed{u^*(x) = \underset{u}{\text{argmin}} \; \mathcal{L}(x,u) + \nabla_x J(x)^T f(x,u)}$$

THIS IS SOLVABLE ONLY FOR THE LINEAR-QUADRATIC CASE, FOR WHICH WE HAVE

$$\dot{P} = 0 \rightsquigarrow \underbrace{0 = Q + PA + A^T P - (S^T + PB) R^{-1} (S + B^T P)}_{\text{THE ALGEBRAIC RICCATI EQUATION IN CONTINUOUS-TIME}}$$

$$\rightsquigarrow u^*(x) = \underbrace{-R^{-1}(S + B^T P)}_{-K} x$$