Dynamic Programming and Optimal Control

Recitation Session

Mohanarajah Gajamohan

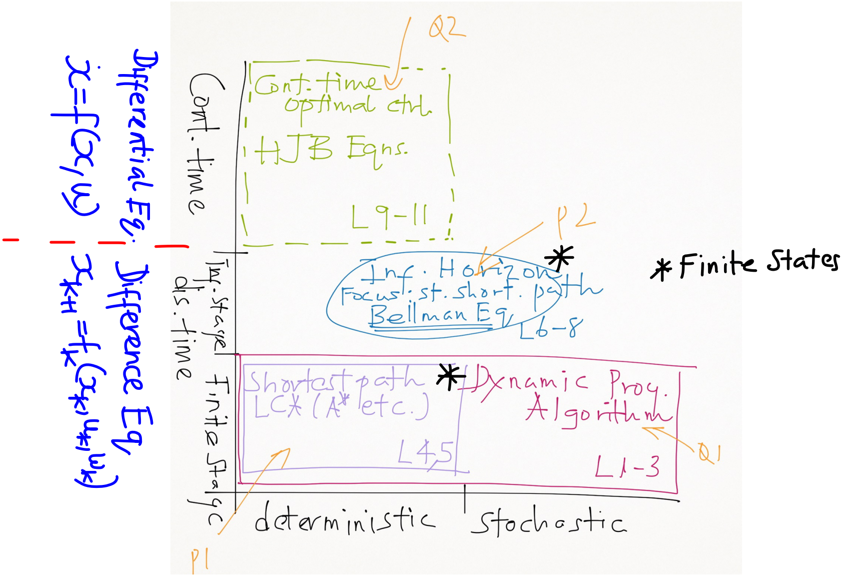
Institute for Dynamic Systems and Control
ETH Zurich, Switzerland
gajan@ethz.ch

December 12, 2012

Plan for Today

1. The Dynamic Programming Algorithm (DPA)
2. Deterministic Systems and the Shortest Path (SP)
3. Infinite Horizon Problems, Stochastic SP
4. Deterministic Continuous-Time Optimal Control

Overview



Q2
Cont. time optimal ctrl.
HJB Eqns.
L9-11
P2
Inf. Horizon
Focus: st. short. path
Bellman Eq.
L6-8
* Finite States
Shortest path
LCA (A* etc.)
L4,5
Dynamic Prog. Algorithm
L1-3
Q1
P1
deterministic
stochastic
Cont. time
Inf. stage
dis. time
Finite stage
Differential Eq.
$\dot{x} = f(x, u)$
Difference Eq.
$x_{k+1} = f_k(x_k, u_k, w_k)$

Dynamic Programming Algorithm

System dynamics:

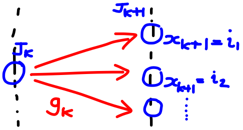
$$x_{k+1} = f_{\underline{k}}(x_k, u_k, w_k), \quad k = 0, 1, \ldots, N-1$$

time varying system

Cost:

$$g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k)$$

DPA:



$$J_N(x_N) = g_N(x_N),$$

$$J_k(x_k) = \min_{u_k \in U_k(x_k)} \mathbb{E}_{w_k} \left\{ g_k(x_k, u_k, w_k) + J_{k+1}(f_k(x_k, u_k, w_k)) \right\}$$

Ex 1: Stage k is not always a time step.

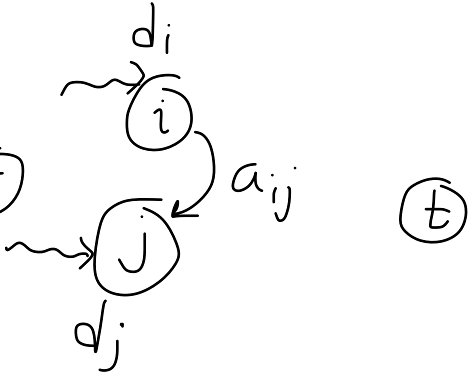
State Augmentation

x_k : summarizes past information that is relevent for future optimization.

Example #2

Shortest Path Problems

Label Correcting Methods



Ex 3: practice at least once!

Note: know the difference between Label correcting variants:
Breadth first, Depth first
pg 84 (Bertsekas)

Shortest Path Problems

Viterbi Algorithm

Andrew Viterbi (1967)

CDMA, GSM, 802.11 (wireless lan)

– decoding convolution code.

Stochastic Shortest Path Problems

Definition $\rho_\pi = \max_{i=1\ldots n} P\{x_m \neq t \mid x_0 = i, \pi\}$

1 2 ... n

$\rho_\pi \leq 1 \ \forall \pi$

t ← termination state

Stochastic Shortest Path Problems

Policy Iteration, Value Iteration, ~~Linear Programming~~

Ex 4:

Continuous time optimal control + Deterministic

System dynamics: $\dot{x}(t) = f(x(t), u(t)), \quad 0 \le t \le T, \quad x(0) = x_0$
Cost: $h(x(T)) + \int_0^T g(x(t), u(t))dt$
HJB:

$$0 = \min_{u \in U} \left[g(x,u) + \frac{\partial V(t,x)}{\partial t} + \left(\frac{\partial V(t,x)}{\partial x} \right)^{\mathrm{T}} f(x,u) \right], \quad \forall x, t \in [0,T],$$

$$V(T,x) = h(x), \quad \forall x.$$

Ex 5:

DPA:

$$J_N(x_N) = g_N(x_N),$$

$$J_k(x_k) = \min_{u_k \in U_k(x_k)} \cancel{\mathbb{E}_{w_k}} \{ g_k(x_k, u_k, \cancel{w_k}) + J_{k+1}(f_k(x_k, u_k, \cancel{w_k})) \}$$

Continuous time optimal control

Minimum principle:

$$H(x, u, p) = g(x, u) + p^{\mathrm{T}} f(x, u)$$

$$\begin{aligned}
\dot{x}^*(t) &= \frac{\partial H(x^*(t), u^*(t), p(t))}{\partial p}, \quad x^*(0) = x_0, \\
\dot{p}(t) &= -\frac{\partial H(x^*(t), u^*(t), p(t))}{\partial x}, \quad p(T) = \frac{\partial h(x^*(T))}{\partial x} \\
u^*(t) &= \operatorname{argmin}_{u \in U} \; H(x^*(t), u(t), p(t))
\end{aligned}$$

Ex 6,7 :