F.L. Lewis
Moncrief-O'Donnell Endowed Chair
Head, Controls & Sensors Group

Supported by :
NSF - PAUL WERBOS
ARO – RANDY ZACHERY

Automation & Robotics Research Institute (ARRI)
The University of Texas at Arlington

Notes on Optimal Control

Draguna Vrabie

The University of Texas ARLINGTON

Talk available online at
http://ARRI.uta.edu/acs

ARRI

Michael K. Sain

Blocking zeros
Zero synthesis
module theory of multivariable zeros of dynamical systems
Exactness of maps
TSP nonlinear feedback synthesis

James L. Massey, Michael K. Sain: Inverse Problems in Coding, Automata, and Continuous Systems FOCS 1967: 226-232

Michael K. Sain: Minimal Torsion Spaces and the Partial Input/Output Problem Information and Control 29(2): 103-124 (1975)

M.K. Sain, B.F. Wyman, J.L. Peczkowski, "Extended zeros and model matching," SIAM Journal on Control and Optimization, May 1991

Cheryl B. Schrader and Michael K. Sain, *"Zero Principles for Implicit Feedback Systems,"* Circuits, Systems, and Signal Processing: Special Issue on Implicit and Robust Systems, Vol. 13, No. 2-3, pp. 273-293, 1994.

Michael K. Sain and Cheryl B. Schrader, *"Feedback, Zeros, and Blocking Dynamics,"* in Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing I. H. Kimura and S. Kodama, eds., Tokyo: Mita Press, pp. 227-232, 1992.

Michael K. Sain and Cheryl B. Schrader, *"Bilinear Operators and Matrices,"* in Mathematics for Circuits and Filters. Wai-Kai Chen, ed., CRC Press, pp. 23-41, 2000.

Ronald W. Diersing, Michael K. Sain, and Chang-Hee Won, "Bi-Cumulant Games: A generalization of H-inf and H_2/H-inf Control," IEEE Transactions on Automatic Control, Submitted, 2007.

Module theoretic zero structures for system matrices

Author(s): Wyman, Bostwick F.; Sain, Michael K.

Abstract: The coordinate-free module-theoretic treatment of transmission zeros for MIMO transfer functions developed by Wyman and Sain (1981) is generalized to include noncontrollable and nonobservable linear dynamical systems. ...

NASA Center: NASA (non Center Specific)
Publication Year: 1987
Added to NTRS: 2004-11-03
Accession Number: 87A30190; Document ID: 19870042916

From the Editor

Michael K. Sain

Editor-in-Chief, IEEE Circuits and Systems Magazine

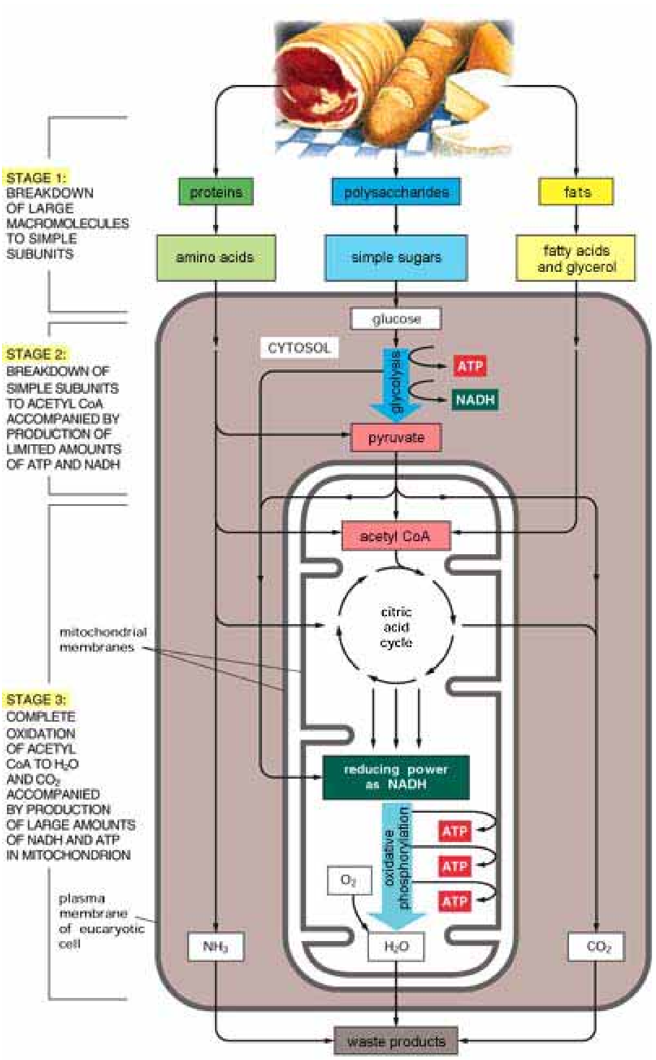
“Did I say that? Well, that was then. This is now.”

- M.K. Sain

EIC editorial, IEEE Circuits & Systems magazine, v.3, no. 1, 2003

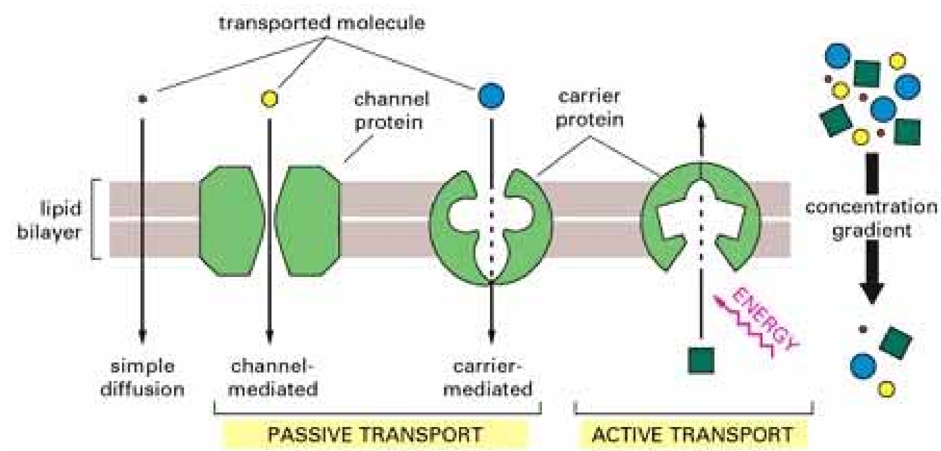
Optimality in Biological Systems

Cell Homeostasis



Cellular Metabolism

The individual cell is a complex feedback control system. It pumps ions across the cell membrane to maintain homeostatis, and has only limited energy to do so.



Permeability control of the cell membrane

http://www.accessexcellence.org/RC/VL/GG/index.html

Optimality in Control Systems Design

R. Kalman 1960

Rocket Orbit Injection

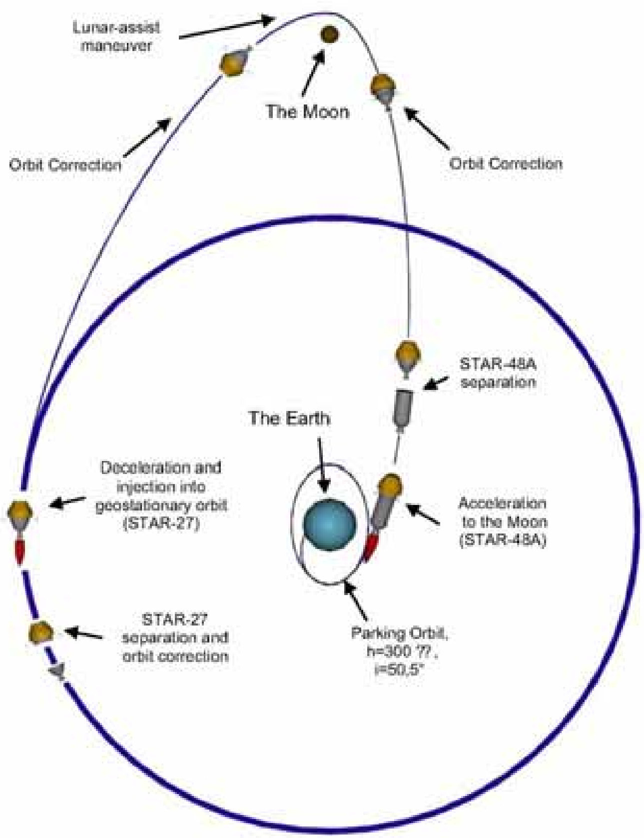


Fig. 1-1. Trajectory scheme

ISC Kosmotras Proprietary 9

http://microsat.sm.bmstu.ru/e-library/Launch/Dnepr_GEO.pdf

Dynamics

$$\dot{r} = w$$

$$\dot{w} = \frac{v^2}{r} - \frac{\mu}{r^2} + \frac{F}{m}\sin\phi$$

$$\dot{v} = \frac{-wv}{r} + \frac{F}{m}\cos\phi$$

$$\dot{m} = -Fm$$

Objectives

- Get to orbit in minimum time
- Use minimum fuel

Adaptive Control is Not Optimal

Optimal Control is off-line,
and needs to know the system dynamics to solve design eqs.

We want ONLINE ADAPTIVE OPTIMAL Control

Continuous-Time Optimal Control

System $\dot{x} = f(x,u)$

Cost $V(x(t)) = \int_t^\infty r(x,u)\,dt = \int_t^\infty (Q(x) + u^T R u)\,dt$

For a given control, the cost satisfies this eq.

Hamiltonian

$$0 = \dot{V} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T \dot{x} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T f(x,u) + r(x,u) \equiv H(x, \frac{\partial V}{\partial x}, u) \quad , \qquad V(0) = 0$$

In LQR, this is a Lyapunov eq

Optimal cost $$0 = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V}{\partial x}\right)^T \dot{x} \right) = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V}{\partial x}\right)^T f(x,u) \right)$$

Bellman $$0 = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T \dot{x} \right) = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T f(x,u) \right)$$

Optimal control $$h^*(x(t)) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V^*}{\partial x}$$

HJB equation $$0 = \left(\frac{dV^*}{dx}\right)^T f + Q(x) - \tfrac{1}{4}\left(\frac{dV^*}{dx}\right)^T g R^{-1} g^T \frac{dV^*}{dx} \quad , \quad V(0) = 0$$

In LQR, this is a Riccati eq

Linear system, quadratic cost

System: $\dot{x} = Ax + Bu$

Utility: $r(x,u) = x^T Qx + u^T Ru;\ R > 0, Q \geq 0$

The cost is quadratic $V(x(t)) = \int_t^{\infty} r(x,u)\,d\tau = x^T(t)Px(t)$

Optimal control (state feedback):

$$u(t) = -R^{-1}B^T Px(t) = -Lx(t)$$

HJB equation is the *algebraic Riccati equation* (ARE):

$$0 = PA + A^T P + Q - PBR^{-1}B^T P$$

Full system dynamics must be known

CT Policy Iteration

To avoid solving HJB equation

Utility $$r(x,u) = Q(x) + u^T Ru$$

Cost for any given u(t)

$$0 = \left(\frac{\partial V}{\partial x}\right)^T f(x,u) + r(x,u) \equiv H(x, \frac{\partial V}{\partial x}, u)$$

Lyapunov equation

Iterative solution

Pick stabilizing initial control

Find cost

$$0 = \left(\frac{\partial V_k}{\partial x}\right)^T f(x, h_k(x)) + r(x, h_k(x))$$

$$V_k(0) = 0$$

Update control

$$h_{k+1}(x) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V_k}{\partial x}$$

- Convergence proved by Saridis 1979 if Lyapunov eq. solved exactly
- Beard & Saridis used complicated Galerkin Integrals to solve Lyapunov eq.
- Abu Khalaf & Lewis used NN to approx. V for nonlinear systems and proved convergence

Full system dynamics must be known

LQR Policy iteration = Kleinman algorithm

1. For a given control policy $u = -L_k x$ solve for the cost:

$$0 = A_k{}^T P_k + P_k A_k + C^T C + L_k{}^T R L_k$$

Lyapunov eq.

$$A_k = A - BL_k$$

2. Improve policy:

$$L_k = R^{-1} B^T P_{k-1}$$

- If started with a stabilizing control policy L_0 the matrix P_k monotonically converges to the unique positive definite solution of the Riccati equation.
- Every iteration step will return a stabilizing controller.
- The system has to be known.

Kleinman 1968

Policy Iteration Solution

Policy iteration

$$(A - BB^T P_i)^T P_{i+1} + P_{i+1}(A - BB^T P_i) + P_i BB^T P_i + Q = 0$$

This is in fact a Newton's Method

$$Ric(P) \equiv A^T P + PA + Q - PBB^T P$$

Then, Policy Iteration is

$$P_{i+1} = P_i - \left(Ric'_{P_i} \right)^{-1} Ric(P_i), \quad i = 0,1,\ldots$$

Frechet Derivative

$$Ric'_{P_i}(P) \equiv (A - BB^T P_i)^T P + P(A - BB^T P_i)$$

Draguna Vrabie

Policy Iterations without Lyapunov Equations

- Dynamic programming
  - built on Bellman’s optimality principle – alternative form for CT Systems [Lewis & Syrmos 1995]

$$V^*(x(t)) = \min_{\substack{u(\tau) \\ t\le\tau<t+\Delta t}} \left\{ \int_t^{t+\Delta t} r(x(\tau),u(\tau))d\tau \quad + \quad V^*(x(t+\Delta t)) \right\}$$

f(x) and *g(x)* do not appear

$$r(x(\tau),u(\tau)) = x^T(\tau)Qx(\tau) + u^T(\tau)Ru(\tau)$$

Draguna Vrabie

Solving for the cost – Our approach

For a given control $u = -Lx$

The cost satisfies

$$V(x(t)) = \int_{t}^{t+T} (x^T Q x + u^T R u)dt \quad + \quad V(x(t+T))$$

f(x) and *g(x)* do not appear

LQR case

$$x(t)^T P x(t) = \int_{t}^{t+T} (x^T Q x + u^T R u)dt + x(t+T)^T P x(t+T)$$

Optimal gain is $L = R^{-1} B^T P$

Draguna Vrabie

Solving for the cost – Our approach

1. Policy iteration

Cost update $$\underline{V_{k+1}}(x(t)) = \int_t^{t+T} (x^T Q x + u^{kT} R u^k) dt + \underline{V_{k+1}}(x(t+T))$$

A and B do not appear

For LQR case

$$u^k(t) = -L_k x(t)$$

$$x^T(t) P_{k+1} x(t) = \int_t^{t+T} x^T(\tau)(Q + L_k{}^T R L_k) x(\tau) d\tau + x^T(t+T) P_{k+1} x(t+T)$$

Control gain update $$L_{k+1} = R^{-1} B^T P_{k+1}$$

B needed for control update

Initial stabilizing control is needed

Lemma 1

$$x^T(t)P_{k+1}x(t) = \int_t^{t+T} x^T(\tau)(Q + L_k{}^T R L_k)x(\tau)d\tau + x^T(t+T)P_{k+1}x(t+T)$$

$$L_k = R^{-1}B^T P_k$$

Solves Lyapunov equation without knowing A or B

is equivalent to

$$(A - BR^{-1}B^T P_k)^T P_{k+1} + P_{k+1}(A - BR^{-1}B^T P_k) + P_k BR^{-1}B^T P_k + Q = 0$$

Proof:

$$\frac{d(x^T P_i x)}{dt} = x^T(A_i{}^T P_i + P_i A_i)x = -x^T(K_i{}^T R K_i + Q)x$$

$$\int_t^{t+T} x^T(Q + K_i{}^T R K_i)x d\tau = -\int_t^{t+T} d(x^T P_i x) = x^T(t)P_i x(t) - x^T(t+T)P_i x(t+T)$$

Theorem

This algorithm converges and is equivalent to Kleinman's Algorithm

Only B is needed

Algorithm Implementation

Critic update

$$x^T(t)P_{k+1}x(t) = \int_t^{t+T} x^T(\tau)(Q + L_k{}^T R L_k)x(\tau)d\tau + x^T(t+T)P_{k+1}x(t+T)$$

Use Kronecker product $vec(ABC) = \left(C^T \otimes A\right)vec(B)$

To set this up as $\bar{x}(t) = x(t) \otimes x(t)$ is the quadratic basis set

$$\bar{p}_{k+1}{}^T \bar{x}(t) = \int_t^{t+T} x(\tau)^T (Q + K_i{}^T R K_i)x(\tau)d\tau + \bar{p}_{k+1}{}^T \bar{x}(t+T)$$

$$\bar{p}_{k+1}{}^T \bar{\varphi}(t) \equiv \quad \bar{p}_{k+1}{}^T \left[\bar{x}(t) - \bar{x}(t+T)\right] = \int_t^{t+T} x(\tau)^T (Q + K_i{}^T R K_i)x(\tau)d\tau$$

$\equiv \rho(t, t+T)$ Reinforcement on time interval *[t, t+T]*

Quadratic regression vector

Now use RLS along the trajectory to get new weights $\bar{p}_{k+1}$

Unpack weights into the matrix P_{k+1}

Then find updated FB gain $L_{k+1} = R^{-1}B^T P_{k+1}$

1. Select initial control policy

2. Find associated cost

Solves Lyapunov eq. without knowing dynamics

$$\bar{p}_{k+1}^{\;T}\left[\bar{x}(t)-\bar{x}(t+T)\right]=\int_{t}^{t+T} x(\tau)^T (Q+L_k^{\;T} R L_k)\, x(\tau)\, d\tau=\rho(t,t+T)$$

3. Improve control $L_{k+1} = R^{-1} B^T P_{k+1}$

observe $x(t)$

apply $u^k(t)=L_k x(t)$

observe cost integral

observe $x(t+T)$

update P

t

t+T

do RLS until convergence to P_{k+1}

update control gain to L_{k+1}

Measure cost increment (reinforcement) by adding V as a state. Then

$$\dot{V} = x^T Q x + (u^k)^T R u^k$$

A is not needed anywhere

Algorithm Implementation

Or use batch Least-Squares solution along the trajectory

The Critic update

$$x^T(t)P_{k+1}x(t)=\int_t^{t+T} x^T(\tau)(Q+L_k^T RL_k)x(\tau)d\tau + x^T(t+T)P_{k+1}x(t+T)$$

can be setup as

$$\bar{p}_{i+1}^T\bar{\varphi}(t) \equiv \bar{p}_{i+1}^T\left[\bar{x}(t)-\bar{x}(t+T)\right]=\int_t^{t+T} x(\tau)^T(Q+L_k^T RL_k)x(\tau)d\tau \equiv d(\bar{x}(t),L_k)$$

$\bar{x}(t) = x(t)\otimes x(t)$ is the quadratic basis set

Evaluating $d(\bar{x}(t),L_k)$ for $n(n+1)/2$ trajectory points, one can setup a least squares problem to solve

$$\bar{p}_{i+1}=(XX^T)^{-1}XY$$

$$X=[\bar{\varphi}^1(t) \quad \bar{\varphi}^2(t) \quad \dots \quad \bar{\varphi}^N(t)]$$

$$Y=[d(\bar{x}^1,K_i) \quad d(\bar{x}^2,K_i) \quad \dots \quad d(\bar{x}^N,K_i)]^T$$

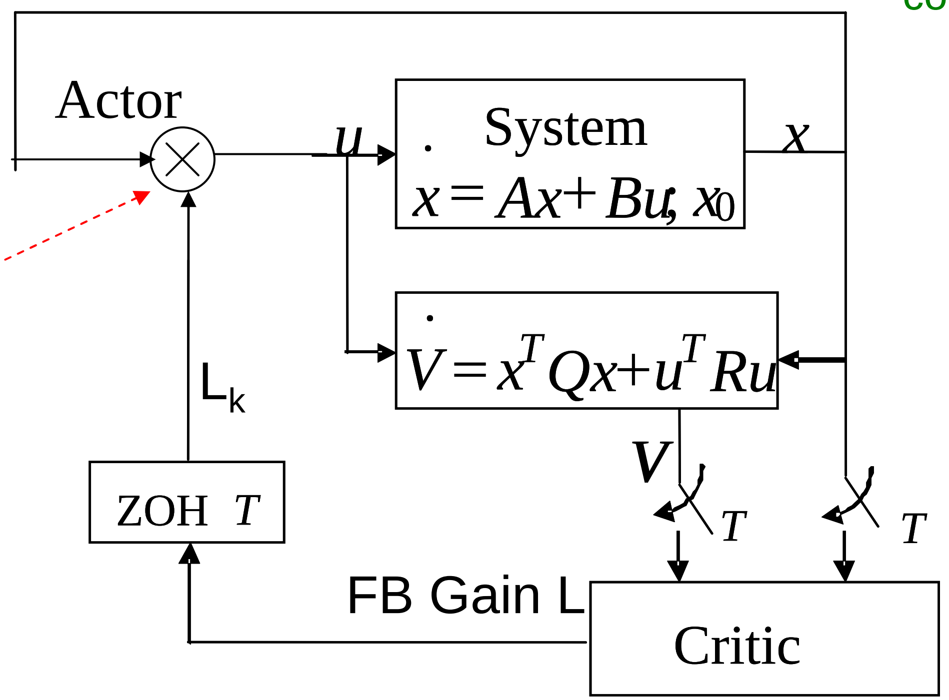
Draguna Vrabie

Direct Optimal Adaptive Controller

"This is a very weird
Control Structure whose likes
I have not seen in
control system theory"

- suspicious quote
by F.L. Lewis, 2007

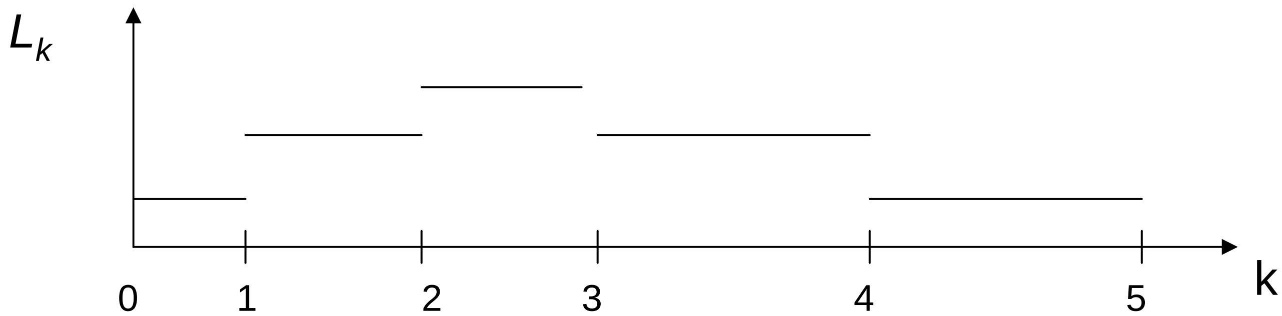
multiplier



Dynamic
Control
System

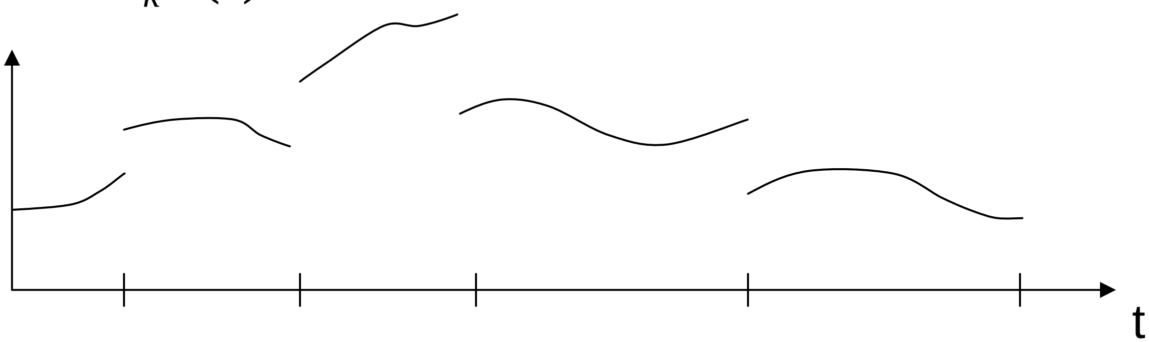
A hybrid continuous/discrete dynamic controller
whose internal state is the observed cost over the interval

Gain update (Policy)



Control

$$u^k(t) = -L_k x(t)$$



Sample periods need not be the same
They can be selected on-line in real time

Continuous-time control with discrete gain updates

Draguna Vrabie

2. CT ADP Greedy iteration

Cost update $$V_{k+1}(x(t)) = \int_t^{t+T} (x^T Q x + u^{kT} R u^k) dt + V_k(x(t+T))$$

Control policy $$u^k(t) = -L_k x(t)$$

LQR $$x^T(t) P_{k+1} x(t) = \int_t^{t+T} x^T(\tau)(Q + L_k{}^T R L_k) x(\tau) d\tau + x^T(t+T) P_k x(t+T)$$

Control gain update

$$L_{k+1} = R^{-1} B^T P_{k+1}$$

A and B do not appear

B needed for control update

No initial stabilizing control needed

Direct Optimal Adaptive Control for Partially Unknown CT Systems

Algorithm Implementation

The critic update

$$x^T(t)P_{k+1}x(t)=\int_t^{t+T} x^T(\tau)(Q+L_k{}^T RL_k)x(\tau)d\tau+x^T(t+T)P_k x(t+T)$$

Use Kronecker product $vec(ABC)=\left(C^T\otimes A\right)vec(B)$

To set this up as $\overline{x}(t)=x(t)\otimes x(t)$ is the quadratic basis set

$$\overline{p}_{k+1}{}^T\overline{x}(t)=\int_t^{t+T} x(\tau)^T(Q+L_k{}^T RL_k)x(\tau)d\tau+\overline{p}_k{}^T\overline{x}(t+T)$$

Regression vector

Previous weights

Now use RLS along the trajectory to get new weights $\overline{p}_{k+1}$

Unpack weights into the matrix P_{k+1}

Then find updated FB gain $L_{k+1}=R^{-1}B^TP_{k+1}$

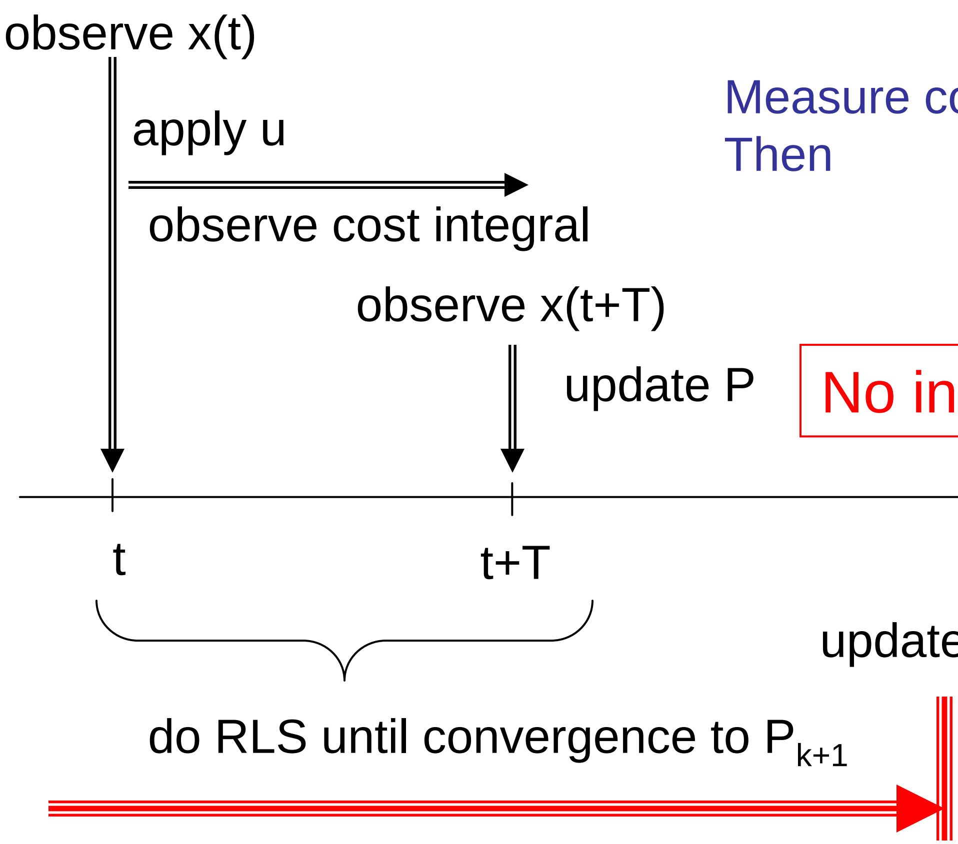
1. Select control policy

2. Find associated cost

Solves for cost update without knowing dynamics

$$\bar{p}_{k+1}^{\;T}\,\bar{x}(t) = \int_{t}^{t+T} x(\tau)^T \left(Q + L_k^{\;T} R L_k\right) x(\tau)\, d\tau + \bar{p}_k^{\;T}\,\bar{x}(t+T)$$

3. Improve control $\quad L_{k+1} = R^{-1} B^T P_{k+1}$



Measure cost increment by adding V as a state.
Then $\quad \dot{V} = x^T Q x + (u^i)^T R u^i$

No initial stabilizing control needed

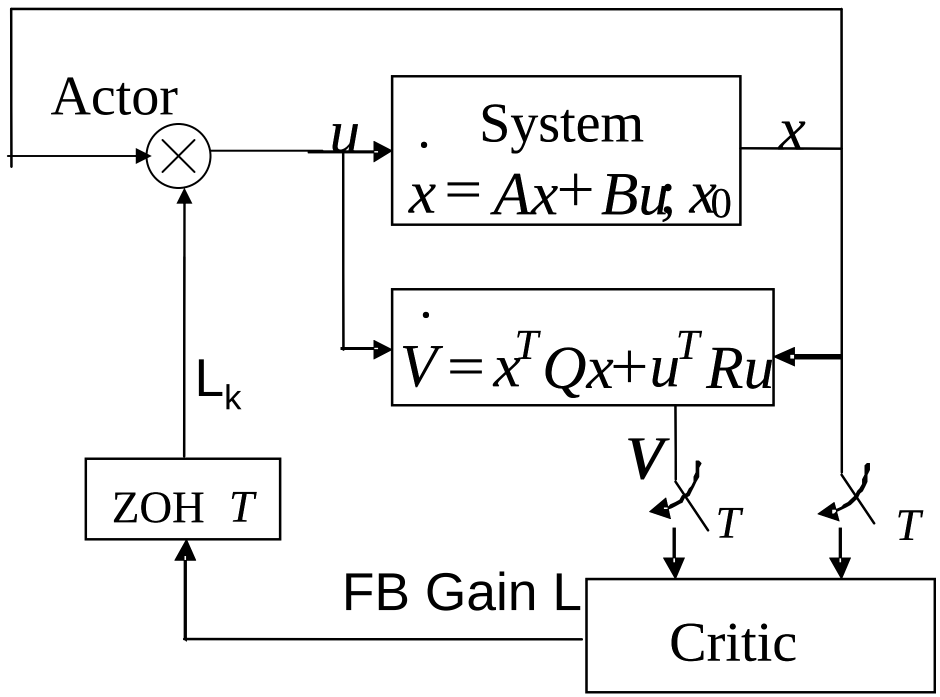
update control gain to L_{k+1}

A is not needed anywhere

Draguna Vrabie

Direct Optimal Adaptive Controller

A hybrid continuous/discrete dynamic controller
whose internal state is the observed value over the interval



Dynamic
Control
System

Has a different critic cost update
No initial stabilizing gain needed

Draguna Vrabie

Analysis of the algorithm

For a given control policy $u^k(t) = -L_k x(t)$ with $L_k = R^{-1} B^T P_k$

$$A_k = A - BR^{-1}B^T P_k$$

Greedy update $V_{k+1}(x(t)) = \int_t^{t+T} \left\{ x^T Q x + \left(u^k\right)^T R u^k \right\} d\tau + V_k(x(t+T)), \quad V_0 = 0$ is equivalent to

$$P_{k+1} = \int_t^{t+T} e^{A_k^T t} (Q + L_k{}^T R L_k) e^{A_k t} dt + e^{A_k^T T} P_k e^{A_k T}$$

a strange pseudo-discretized RE

c.f. DT RE

$$P_{k+1} = \overline{A}^T P_k \overline{A} + Q - \overline{A}^T P_k B \left(P_k + B^T P_K B \right)^{-1} B^T P_k \overline{A}$$

$$P_{k+1} = \overline{A}_k^T P_k \overline{A}_k + Q + L_k^T \left(P_k + B^T P_K B \right) L_k$$

Draguna Vrabie

Analysis of the algorithm

Lemma 1. The ADP iteration between (13) and (14) is equivalent to the Quasi-Newton method

$$P_{i+1} = P_i - (Ric'_{P_i})^{-1}\left(Ric(P_i) - e^{A_i T^T} Ric(P_i) e^{A_i T} \right). \quad (19)$$

This extra term means the initial Control action need not be stabilizing

Lemma 2. CT HDP is equivalent to

$$P_{k+1} - P_k = \int_0^T e^{A_k^T t}(P_k A + AP_k + Q - L_k^T RL_k)e^{A_k t} dt \qquad A_k = A - BR^{-1}B^T P_k$$

When ADP converges, the resulting P satisfies the Continuous-Time ARE !!

Lemma 3. Let the ADP algorithm converge so that $P_i \to P^*$. Then P^* satisfies $Ric(P^*) = 0$, i.e. P^* is the solution the continuous-time ARE.

ADP solves the CT ARE without knowledge of the system dynamics A

Solve the Riccati Equation
WITHOUT knowing the plant dynamics

Model-free ADP

Direct OPTIMAL ADAPTIVE CONTROL

Works for Nonlinear Systems

a neural network is used to approximate the cost

Robustness?
Comparison with adaptive control methods?

Policy Evaluation – *Critic update*

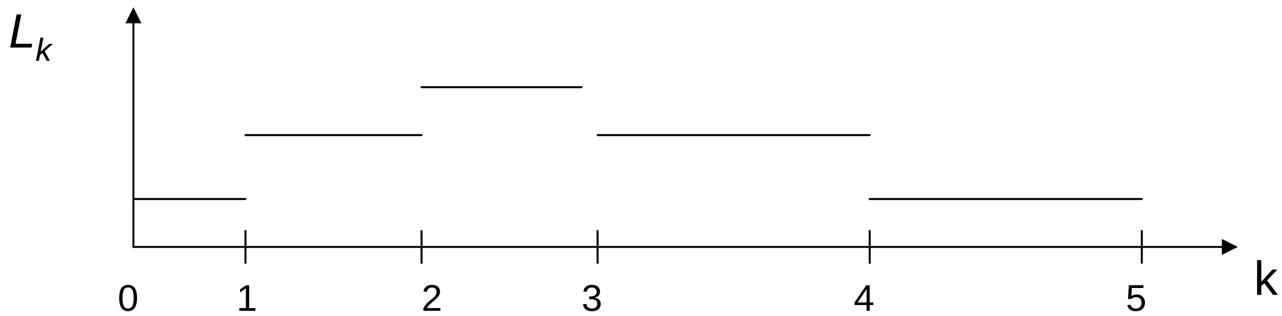
Let K be *any* state feedback gain for the system (1). One can measure the associated cost over the infinite time horizon

$$V(t,x(t)) = \int_t^{t+T} x(\tau)^T (Q + K^T R K) x(\tau) d\tau + \underbrace{W(t+T, x(t+T))}$$

where $W(t+T, x(t+T))$ is an initial infinite horizon cost to go.

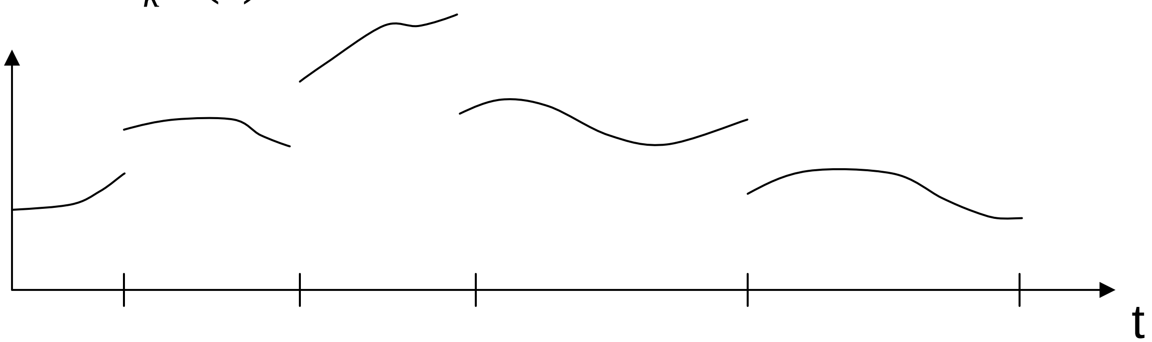
What to do about the tail – issues in Receding Horizon Control

Gain update (Policy)



Control

$$u^k(t) = -L_k x(t)$$



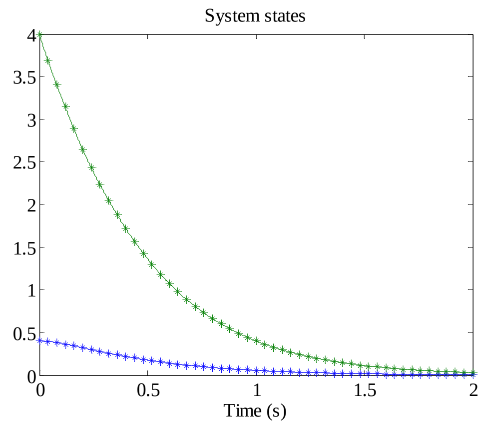
Sample periods need not be the same

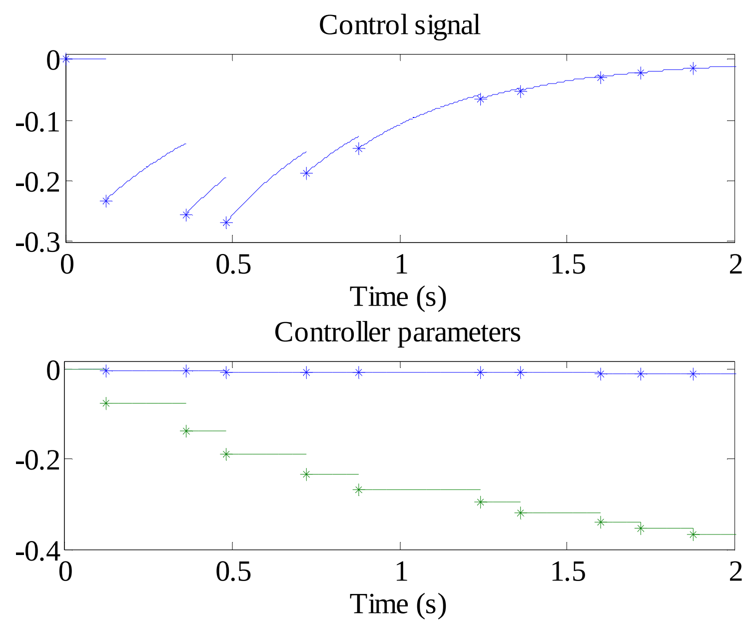
Continuous-time control with discrete gain updates

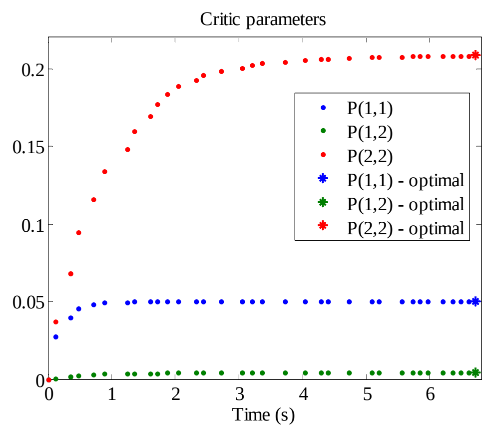
Simulations on: F-16 autopilot

Load frequency control for power system

A matrix not needed







Converge to SS Riccati equation soln

Continuous-Time Optimal Control

System $\dot{x} = f(x,u)$

c.f. DT value recursion, where f(), g() do not appear

Cost $V(x(t)) = \int_t^\infty r(x,u)\,dt = \int_t^\infty (Q(x) + u^T R u)\,dt$

$$V_h(x_k) = r(x_k, h(x_k)) + \gamma V_h(x_{k+1})$$

Hamiltonian

$$0 = \dot{V} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T \dot{x} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T f(x,u) + r(x,u) \equiv H(x, \frac{\partial V}{\partial x}, u) \qquad V(0) = 0$$

Optimal cost

$$0 = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V}{\partial x}\right)^T \dot{x} \right) = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V}{\partial x}\right)^T f(x,u) \right)$$

Bellman

$$0 = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T \dot{x} \right) = \min_{u(t)} \left(r(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T f(x,u) \right)$$

Optimal control

$$h^*(x(t)) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V^*}{\partial x}$$

HJB equation

$$0 = \left(\frac{dV^*}{dx}\right)^T f + Q(x) - \tfrac{1}{4}\left(\frac{dV^*}{dx}\right)^T g R^{-1} g^T \frac{dV^*}{dx} \qquad V(0) = 0$$