Ottimizzazione nei Sistemi di Controllo 1
a.a. 2020/2021

Mario Sassano

Dipartimento di Ingegneria Civile e Ingegneria Informatica
Università di Roma Tor Vergata

Informazioni sul corso

E-mail: mario.sassano@uniroma2.it / mario.sassano08@gmail.com

Sito Web: http://control.disp.uniroma2.it/sassano/Teaching.html

Orario: Martedì 14:00-15:45 Aula C6, Mercoledì 16:00-17:45 Aula C2

Riferimenti e libri consigliati:

- O.M. Grasselli, L. Menini e S. Galeani, *Sistemi Dinamici - Introduzione all'analisi e primi strumenti di controllo*
- R.S. Sutton e A. G. Barto, *Reinforcement Learning: An introduction*
  $\Rightarrow$ web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf
- P. Dorato, C.T. Abdallah e V. Cerone, *Linear Quadratic Control - An introduction*
- D. Liberzon, *Calculus of Variations and Optimal Control Theory - A concise introduction*
  $\Rightarrow$ liberzon.csl.illinois.edu/teaching/cvoc.pdf
- J. Engwerda, *LQ Dynamic Optimization and Differential Games*

Introduzione alla ottimizzazione dinamica (1/3)

Ottimizzazione dinamica

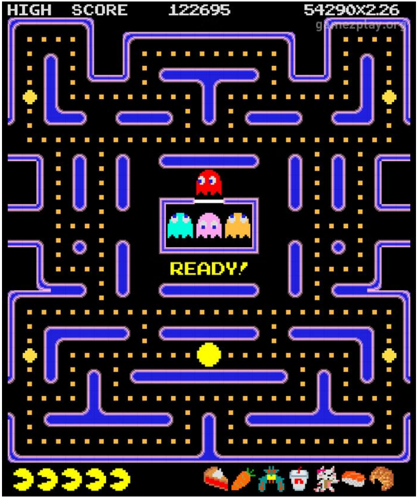
Problema di **minimizzazione/massimizzazione** in cui le variabili decisionali e/o altri parametri che definiscono il problema potenzialmente **variano** nel tempo

- per **cause esterne** (cambiamento del contesto, ad esempio comprare una azione...);
- in **conseguenza di scelte precedenti**, in modo
  - non istantaneo
  - non banale

Ottimizzazione dinamica è caratterizzata da

- **funzione obiettivo cumulativa** di costi/ricavi istantanei
- **vincoli dinamici**

Esempio di problema di ottimizzazione dinamica:



- **Vincolo dinamico** definito dall'evoluzione della posizione di PacMan nel labirinto
- **Sequenza di decisioni** da prendere (destra, sinistro, dritto)
- **Ricompense raccolte** di volta in volta durante l'evoluzione
- Obiettivo è **massimizzare** la somma di tutte le ricompense
- **Soluzione statica impossibile** per la presenza degli inseguitori

Ottimizzazione dinamica

Problema di **minimizzazione/massimizzazione** in cui le variabili decisionali e/o altri parametri che definiscono il problema potenzialmente **variano** nel tempo

- per **cause esterne** (cambiamento del contesto, ad esempio comprare una azione...);
- in **conseguenza di scelte precedenti**, in modo
  - non istantaneo
  - non banale

Ottimizzazione dinamica è caratterizzata da

- **funzione obiettivo cumulativa** di costi/ricavi istantanei
- **vincoli dinamici**

In casi semplici

$\Downarrow$

Soluzione **statica con discretizzazione temporale** su orizzonte temporale fissato e finito..

Lot-sizing: decidere quando e quanto produrre per soddisfare una domanda (nota)

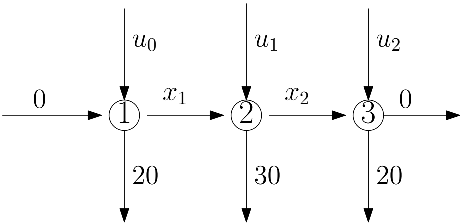
- costi fissi di produzione ⇒ pochi lotti grandi
- costi di immagazzinamento ⇒ lotti piccoli "quando serve"

Supponiamo $N = 3$

$$\begin{cases} \min\limits_{u_0, u_1, u_2} & \sum_{k=0}^{N}(p_k u_k + A_k \eta(u_k) + h_k x_k) \\ & 0 \leqslant u_k \leqslant M, \quad k = 0,1,2 \\ & 0 \leqslant x_k, \quad k = 1,2,3 \\ & x_1 = u_0 - 20 \\ & x_2 = x_1 + u_1 - 30 \\ & 0 = x_2 + u_2 - 20 \end{cases}$$

con

$$\eta(u_k) = \begin{cases} 0, & u_k = 0 \\ 1, & u_k > 0 \end{cases}$$



- p_k: costo di produzione al tempo k
- h_k: costo di immagazzinamento
- A_k: costo fisso di produzione
- M: limite di produzione

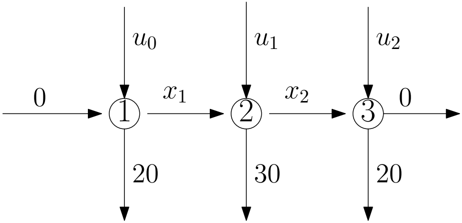
Programmazione lotti di produzione (2/3)

Lot-sizing: decidere quando e quanto produrre per soddisfare una domanda (nota)

- costi fissi di produzione ⇒ pochi lotti grandi
- costi di immagazzinamento ⇒ lotti piccoli "quando serve"

Supponiamo $N = 3$

$$\begin{cases} \min\limits_{u_0, u_1, u_2} & \sum_{k=0}^{N} (p_k u_k + A_k \eta(u_k) + h_k x_k) \\ & 0 \leqslant u_k \leqslant M, \quad k = 0, 1, 2 \\ & 0 \leqslant x_k, \quad k = 1, 2, 3 \\ & \color{red}{x_1 = u_0 - 20} \\ & \color{red}{x_2 = x_1 + u_1 - 30} \\ & \color{red}{0 = x_2 + u_2 - 20} \end{cases}$$



I **vincoli** in rosso riassumono le conseguenze delle scelte precedenti in modo **esplicito**!

⇒ se N cresce (potenzialmente, $N \to \infty$...) diventa impossibile computazionalmente

I **vincoli** in rosso hanno una "natura" diversa da vincoli algebrici classici...

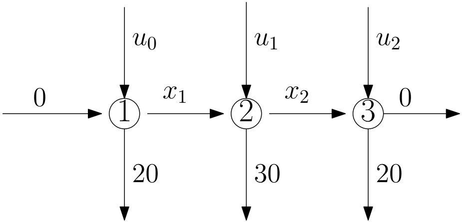
⇒ **Sistema dinamico**!

Lot-sizing: decidere quando e quanto produrre per soddisfare una domanda (nota)

- costi fissi di produzione ⇒ pochi lotti grandi
- costi di immagazzinamento ⇒ lotti piccoli "quando serve"

Supponiamo $N = 3$

$$\begin{cases} \min\limits_{u_0, u_1, u_2} & \sum_{k=0}^{N} (p_k u_k + A_k \eta(u_k) + h_k x_k) \\ & 0 \leqslant u_k \leqslant M, \quad k = 0, 1, 2 \\ & 0 \leqslant x_k, \quad k = 1, 2, 3 \\ & \\ & x_{k+1} = x_k + u_k - d_k, x_0 = 0, \quad k \in \mathbb{N} \end{cases}$$



Evoluzione temporale delle variabili coinvolte deve essere rappresentata in modo **implicito** attraverso l'utilizzo di modelli dinamici

Commenti

- Orizzonte temporale arbitrariamente lungo, ma anche *arbitrariamente fitto*

  $\Rightarrow$ utilizzo di modelli dinamici (lineari) a **tempo continuo**

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \in \mathbb{R}$$

  $\rightarrow$ soluzione più compatta

  $\rightarrow$ soluzione "a posteriori" rispetto a campionamenti arbitrari

- *Infinito è soggettivo* (dipende dalle costanti di tempo del modello dinamico)

  $\Rightarrow$ ottimizzazione regime permanente vs evoluzione transitoria

- *Soluzioni open-loop vs closed-loop*

  $\Rightarrow$ tecniche di ottimizzazione dinamica forniscono la soluzione in funzione dello stato corrente x_k (strategia o *policy*)

Studiamo

1) tecniche di **Ottimizzazione dinamica** (Controllo ottimo/ Programmazione Dinamica, Model Predictive Control, Reinforcement learning, Teoria dei giochi)
⇒ selezione di leggi di controllo in retro-azione (policy) per minimizzare un costo (o massimizzare un ricavo) cumulativo

in presenza di sistemi lineari/nonlineari stazionari[1]

$$\dot{x}(t) = Ax(t) + Bu(t)/\dot{x}(t) = f(x(t), u(t)) \quad t \in \mathbb{R}$$

$$x_{k+1} = Ax_k + Bu_k/x_{k+1} = f(x_k, u_k) \quad k \in \mathbb{N}$$

dove $x \in \mathbb{R}^n$ è lo stato del sistema e $u \in \mathbb{R}^m$ è il controllo attraverso cui agire

imponendo (se $N = +\infty$) la proprietà di

2) **stabilità**[2] **asintotica**[3], per garantire che scelte passate non portino ad *inconsistenze* nel futuro

[1] La soluzione di un sistema stazionario non dipende dall'istante iniziale

[2] piccole variazioni sulla condizione iniziale di equilibrio comportano piccole perturbazioni della soluzione corrispondente..

[3] .. le traiettorie perturbate convergono all'equilibrio per t che tende ad infinito

Esempio: supponiamo di dover spostare su un piano una massa m da una posizione iniziale X_i ad una posizione desiderata X_f imprimendo una forza F

Dal secondo principio della dinamica ($F = ma$, a accelerazione) e definendo $x_1 = X$ (posizione della massa), $x_2 = \dot{X}$ (velocità della massa) e $u = F$ (forza impressa), l'evoluzione temporale $x(t) = (x_1(t), x_2(t))$ è descritta da

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} u \triangleq Ax + Bu$$

⇒ legge di **controllo open-loop** (da Teoria dei Sistemi, con ipotesi di controllabilità[4])

$$u_{ol}(t) = -B^\top e^{A^\top (T-t)} \left(\int_0^T e^{At} B B^\top e^{A^\top t} dt \right)^{-1} e^{AT} \begin{bmatrix} X_i \\ 0 \end{bmatrix} = -6(1-2t)X_i$$

tale che $x_1(0) = X_i$ e $x_1(T) = X_f = 0, \ T = 1s$

☹ **poco robusta**

⇒ dipende solo da X_i, A e B e non da $x(t)$

☺ legge di controllo a **minimia energia**

⇒ $\min_{u(\cdot)} \int_0^T u(t)^\top u(t) dt$ tra tutte le leggi che trasferiscono lo stato da X_i a X_f

[4] Il sistema si dice controllabile se per ogni condizione iniziale $\bar{x}$ esistono un tempo finito $\bar{t}$ e una funzione del tempo $\bar{u}(\cdot)$ tali che $x(0) = \bar{x}$ e $x(\bar{t}) = 0$

Obiettivo: estendere la soluzione precedente in due aspetti

(1) legge di controllo da determinare nella classe di controllori in **retro-azione dallo stato** $u(t) = Kx(t)$ (nel caso lineare), con matrice $K \in \mathbb{R}^{m\times n}$ da determinare o $u(t) = \pi(x(t))$ (nel caso nonlineare)

(2) **funzionali**[5] **di costo** più generali

$$J(u) = \int_0^T \ell(x(t), u(t))dt$$

tenendo in conto sia lo sforzo di controllo $u(t)$ (minima energia) che il valore dello stato corrente $x(t)$ ("risultato" ottenuto)

[5] Funzionale è una funzione a valori in $\mathbb{R}$, $J(\cdot)$, di una funzione, $u(\cdot)$

Programmazione Dinamica (DP)

☺☺☺ fornisce condizioni necessarie e sufficienti di ottimalità

☺☺ condizioni ricorsive per sistemi dinamici a tempo discreto

☹ equazione alle derivate parziali, equazioni di Hamilton-Jacobi-Bellman, per sistemi nonlineari a tempo continuo

☺ equazione matriciale quadratica, equazione di Riccati, per sistemi lineari a tempo continuo e con $\ell(x, u) = x^\top Q x + u^\top R u$

Applicazioni del controllo ottimo

Perchè Controllo Ottimo e DP?

- Dynamic Pricing, Revenue Management
- Ottimizzazione di investimenti
- Gestione di reti elettriche (Smart Grids)
- Auto a guida autonoma
- Applicazioni aerospaziali
- Osservazione della natura (formazione di cristalli, ...)

Applicazioni del controllo ottimo

Perchè Controllo Ottimo e DP?

- Dynamic Pricing, Revenue Management
  $\Rightarrow$ Scegliere il prezzo giusto al momento giusto per massimizzare i ricavi

Lufthansa Systems

SOLUTIONS / COMMERCIAL SOLUTIONS / REVENUE MANAGEMENT & PRICING

Revenue Management & Pricing

We offer fare management and revenue management solutions and a solution to improve your route profitability.

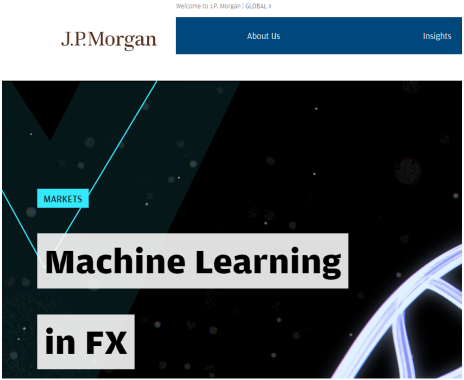
Revenue management and pricing, the cornerstones of a commercial strategy

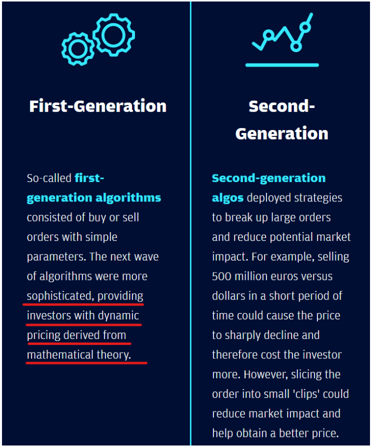
Today's airline business is evolving into a two-tier industry: global alliances are reaching worldwide coverage and no-frills carriers are gaining market share with a low-cost, point-to-point product. Yet both sides recognize revenue management and pricing as cornerstones of their commercial strategy. Sophisticated systems have been designed that optimally support these processes. The future will bring a closer integration of revenue management and pricing systems. Imagine the benefit of forecasts that trigger automatic fare changes, or using price-elasticity models to maximize revenue automatically on specific segments.

Applicazioni del controllo ottimo

Perchè Controllo Ottimo e DP?

- Ottimizzazione di investimenti





Applicazioni del controllo ottimo

Perchè Controllo Ottimo e DP?

- Applicazioni aerospaziali

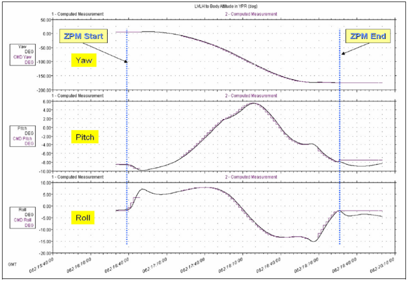
Manovra a carburante nullo per la Stazione Spaziale Internazionale tramite utilizzo dinamico di *Momentum storage devices*

ZERO PROPELLANT MANEUVER™ FLIGHT RESULTS FOR 180° ISS ROTATION

Nazareth Bedrossian[1], Sagar Bhatt[2]
The Charles Stark Draper Laboratory, Inc., Houston, TX 77058

Mike Lammers[3], Louis Nguyen[4]
NASA Johnson Space Center, Houston, TX, 77058

This paper presents results for the Zero Propellant Maneuver (ZPM) ™ attitude control concept flight demonstration. On March 3, 2007, a ZPM was used to reorient the International Space Station 180 degrees without using any propellant. The identical reorientation performed with thrusters would have burned 110lbs of propellant. The ZPM was a pre-planned trajectory used to command the CMG attitude hold controller to perform the maneuver between specified initial and final states while maintaining the CMGs within their operational limits. The trajectory was obtained from a PseudoSpectral solution to a new optimal attitude control problem. The flight test established the breakthrough capability to simultaneously perform a large angle attitude maneuver and momentum desaturation without the need to use thrusters. The flight implementation did not require any modifications to flight software. This approach is applicable to any spacecraft that are controlled by momentum storage devices.



Perchè Controllo Ottimo e DP?

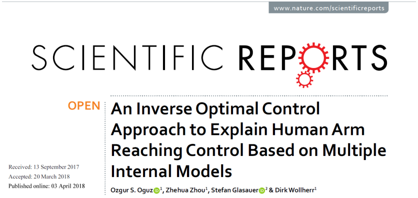
- Osservazione della natura
  ⇒ Fratturazione a minima energia e massima superficie

Applicazioni del controllo ottimo

Perchè Controllo Ottimo e DP?

- Controllo ottimo inverso in fisiologia

  $\Rightarrow$ compromesso ottimo tra cinematica e dinamica del movimento


www.nature.com/scientificreports

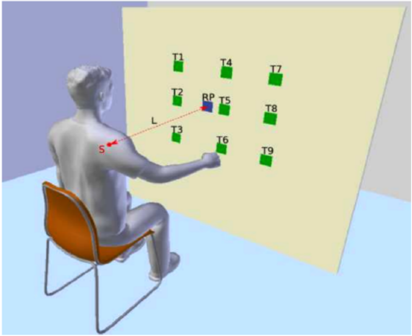
SCIENTIFIC REPORTS

OPEN

Received: 13 September 2017
Accepted: 20 March 2018
Published online: 03 April 2018

An Inverse Optimal Control Approach to Explain Human Arm Reaching Control Based on Multiple Internal Models

Ozgur S. Oguz[1], Zhehua Zhou[1], Stefan Glasauer[2] & Dirk Wollherr[1]



La Programmazione Dinamica si basa sulla conoscenza esatta del modello

In assenza di un modello esatto è possibile **imparare** dall'esperienza (dati, costi o ricompense, raccolti) a seguito di azioni eseguite?

$\Rightarrow$ **Reinforcement learning** (RL)

- determinare il *valore* di ciascuno stato (Value iteration)
- determinare e migliorare il valore di ciascuna azione (Policy iteration)
- implementare algoritmi di DP esclusivamente basati su dati raccolti (Q-learning)

Perchè RL?

- algoritmi di apprendimento *goal-oriented*
- intelligenza artificiale

Ottimizzazione dinamica multi-obiettivo

- definizione: N giocatori *condividono* lo stato di un sistema dinamico

$$\dot{x}(t) = Ax(t) + B_1 u_1(t) + \dots + B_N u_N(t)$$

  e devono individualmente prendere decisioni in modo da ottimizzare un proprio funzionale di costo

$$J_i(u_1, \dots, u_N) = \int_0^T \left(x(t)^\top Q_i x(t) + \sum_{j=1}^N u_j(t)^\top R_{ij} u_j(t)\right) dt$$

  (potenzialmente in conflitto con gli altri)

- concetto di soluzione: **Equilibrio di Nash**
  $\Rightarrow$ non si può migliorare deviando *unilateralmente* dalla propria strategia

- soluzione costruita a partire dalle soluzioni di problemi di controllo ottimo per i singoli giocatori
  $\Rightarrow$ equazioni di Riccati accoppiate