Competitive Algorithms for Online Control

Yisong Yue
This Talk

• Bulk of work in collaboration with Adam Wierman’s group.
Online Optimization

$\mathbf{f}$

$x$

$c(x_1, x_0)$

$\mathbf{f}(x)$

$f_1(x_1)$

$x_0$

$x_1$

Material from Adam Wierman
Online Optimization

Material from Yiheng Lin
Online Optimization

Material from Yiheng Lin
Primer: Smoothed Online Convex Optimization (SOCO)

Goal: Design algorithms to minimize the total cost.

Performance metric: competitive ratio: $\sup \{f_t\} \frac{\text{cost(ALG)}}{\text{cost(OPT)}}$
Comments on Competitive Ratio

\[ \sup_{f_t} \frac{\text{cost}(ALG)}{\text{cost}(OPT)} \]

- Stronger criteria than regret
- Offline OPT gets to see the future
- Goal is typically to achieve constant competitive ratio
  - (note that CCR typically leads to linear regret)
Example 1: Sustainable Data Centers

Can a data center run (almost) entirely on renewable sources?

Requires **dynamic rightsizing** of capacity and **smart deferral** of workloads

*Material from Adam Wierman*
Example 1: Sustainable Data Centers

Can a data center run (almost) entirely on renewable sources?
Example 2: Tracking Control under Predictable Disturbances
Connection to Convex Body Chasing

\[ \|x_2 - x_1\| \]

Material from Adam Wierman
Connection to Convex Body Chasing

Material from Adam Wierman
How do you decide where to move without knowing the future?

Connection to Convex Body Chasing

Material from Adam Wierman
Clearly, a SOCO algorithm can solve CBC...

...a CBC algorithm can also solve a SOCO instance!

Online Control

• (Discrete Time) Dynamical System: \( x_{t+1} = g(x_t, u_t, w_t) \)

  Linear Time Invariant: \( g(x_t, u_t, w_t) = Ax_t + Bu_t + w_t \)

• Additive Control Objective: \( \sum_t \text{Cost}_t(x_t, u_t) \)

  Quadratic: \( \text{Cost}_t(x_t, u_t) = x_t^T Q x_t + u_t^T R u_t \)

• LQR: LTI System w/ Additive Quadratic Control Cost
Online Control (cont.)

- Restrict to LQR-like settings:
  \[
  \min \sum_{t} x_t^T Q x_t + u_t^T R u_t \\
  \text{s.t. } x_{t+1} = A x_t + B u_t + w_t
  \]

- Problem Settings:
  - Do we know A & B a priori or must we learn?
  - What assumptions can we make on A & B?
  - Is \( x_t \) fully or partially observed?
  - Is \( w_t \) observed before or after committing to \( u_t \)?
  - What assumptions can we make on \( w_t \)?
  - How do we measure performance?

- Known A & B, stabilizable
- Fully observed
- \( w_{t:t+\ell-1} \) predictable & bounded
- Competitive Ratio
Recall: Tracking Control under Predictable Disturbances
Competitive Control (for LQR-like problems)

At each time $t$:
- Observe $x_t, \hat{w}_t, \ldots, \hat{w}_{t+\ell-1}$
- Choose $u_t$
- Repeat

Minimize Competitive Ratio:

$$
\sup_{x_0, w_1, \ldots, w_T} \frac{\text{cost(ALG)}}{\text{cost(OPT)}}
$$

---

(State) $x_t \xrightarrow{T} Q x_t + u_t R u_t$

(Control) $x_{t+1} = Ax_t + Bu_t + w_t$

(Disturbance) (imperfect) knowledge of $w_t, \ldots, w_{t+\ell-1}$

“Fixed-Horizon Control” or “Model-Predictive Control”

Focus on 1-step prediction: $\ell = 1$
Comments on Best Linear Controller & Static Regret

- Common to measure regret v.s. best static linear controller:
  - \( u_t = K(x_t^* - x_t) \)

- However, the best static linear controller may have arbitrarily large competitive ratio vs offline optimal.

Reduction to Online Convex Optimization w/ Structured Memory

• Special case of LQR: Input-Disturbed Squared Regulator (IDSР):
  • (finalizing results on general LQR setting)

\[
\begin{align*}
\min \sum_t q_t \frac{1}{2} ||x_t||^2 + \frac{1}{2} ||u_t||^2 \\
s.t. x_{t+1} = Ax_t + B(u_t + w_t)
\end{align*}
\]

• Many robotic systems are well described by IDSРs
  • E.g., https://arxiv.org/abs/1811.08027

Online Optimization with Memory and Competitive Control, Guanya Shi, Yiheng Lin, et al., NeurIPS 2020
Reduction to Online Convex Optimization w/ Structured Memory

• OCO w/ Structured Memory

\[
\begin{align*}
\min & \sum_t q_t \|x_t\|^2 + \frac{1}{2} \|u_t\|^2 \\
\text{s.t.} & \quad x_{t+1} = Ax_t + B(u_t + w_t)
\end{align*}
\]

\[
\min \sum_t f_t(y_t) + \frac{1}{2} \left\| y_t - \sum_{i=1}^{p} C_i y_{t-i} \right\|^2
\]

• \(y_t\) is a transformed representation of state \(x_t\)
  • \(p\) depends on dynamics (single, double integrator, etc.)
• Choose next state indirectly via control action \(u_t\)
• Knowing \(w_t\) defines OCO hitting & switching costs

New OCO setting!

Online Optimization with Memory and Competitive Control, Guanya Shi, Yiheng Lin, et al., NeurIPS 2020
Roadmap to Optimistic ROBD

• Online Balanced Descent (OBD)

• Regularized Online Balanced Descent (ROBD)

• Optimistic ROBD
Online Balanced Descent (OBD)


Choose level set to balance hitting and switching costs.
Greedy/Regularized Online Balanced Descent


Equivalent to regularization: \( \arg\min_y f_t(y) + \lambda_1 c(y, y_{t-p:t-1}) + \frac{\lambda_2}{2} \|y - v_t\|^2 \)
Optimistic ROBD

\[ w_t \in \Omega_t \ (w_t \text{ approximately known}) \rightarrow f_t \text{ approximately known} \]

Choose \( w_t \) optimistically to minimize total cost.
Optimistic ROBD

\[ w_t \in \Omega_t \ (w_t \text{ approximately known}) \rightarrow f_t \text{ approximately known} \]

Choose \( w_t \) optimistically to minimize total cost

(Simplified) CR:

\[ O\left( (q_{max} + 4\alpha^2) \max \left\{ \frac{1}{\lambda}, \frac{\lambda + q_{min}}{(1 - \alpha^2)\lambda + q_{min}} \right\} \right) \]
Summary

• Competitive ratio is a strong benchmark (practically relevant)

• New algorithms for competitive control
  • Online Optimization w/ Structured Memory

• Also working on forecasting & delay

• Also working on nonlinear dynamical systems
  • Online Learning of Nonlinear Control with Guaranteed Success: An Oracle-Based Robust Control-Powered Approach *(preprint coming soon)*
Online Optimization with Memory and Competitive Control, Guanya Shi, Yiheng Lin, et al., NeurIPS 2020

On the Power of Predictions in Online Control, Chenkai Yu, Guanya Shi, et al., NeurIPS 2020

Competitive Control with Delayed Imperfect Information, Chenkai Yu, Guanya Shi, et al., arXiv

Online Learning of Nonlinear Control with Guaranteed Success: An Oracle-Based Robust Control-Powered Approach, Dimitar Ho, Hoang Le, et al., (preprint available soon)