AGNOSTIC CONTROL THEORY

cf

JOINT WORK WITH C. W. ROWLEY

and a team of present & former

GRAD STUDENTS & POSTDOCS.

The Team

J. Carruth

M. Eggl

D. Goswami

B. Guillen

D. Gurevich

A. Simha

M. Weber

Supported by AFOSR

AGNOSTIC CONTROL THEORY

PERTAINS TO CONTROL WITH

LEARNING ON THE FLY,

EG. LANDING AN AIRPLANE AFTER A BIRDSTRIKE HAS KILLED ALL ITS ENGINES, OR AFTER A MISSILE HAS SHOT OFF ONE WING.

"ADAPTIVE CONTROL THEORY"

TRIES TO FIND STRATEGIES WHOSE

PENALTY AFTER TIME T

HAS FAVORABLE ASYMPTOTICS AS $T \to \infty$.

WE WILL WANT TO MINIMIZE A PENALTY

FOR GIVEN, FIXED T

(BUT WE LOOK ONLY AT SIMPLE 1D PROBLEMS)

PLAN OF THE TALK:

Will present a PDE.

The team can solve it numerically

in a relevant range of parameters,

but I know NO RIGOROUS RESULTS

on that equation.

This talk will explain what would follow
if we knew that the PDE admits solutions.

WE ENCOURAGE ANYONE INTERESTED
to SOLVE THE PDE FOR US!

We will:

- Present the PDE

- Present Classical & Nonclassical Problems of Control Theory.

- State Thms Proven Modulo an assumption That the PDE admits a Reasonable Solution.

Let's Get Started ---

THE PDE

In our PDE, the unknown function S

is

$$S(q, t, \xi_1, \xi_2)$$

where

$t \in [0, T]$ (Time)

$q \in \mathbb{R}$ (position),

$\left.\begin{array}{l} \xi_2 \in (0, \infty) \\ \xi_1 \in \mathbb{R} \end{array}\right\}$ (we'll see)

THE EQUATION :

$$\frac{1}{4}(\partial_q S)^2 - q^2 = \left[\partial_t + \bar{a} q \left(\partial_q + q\,\partial_{\xi_1}\right) + q^2 \partial_{\xi_2} + \frac{1}{2}\left(\partial_q + q\,\partial_{\xi_1}\right)^2\right] S$$

WITH

$$S\big|_{t=T} = 0 .$$

$$\bar{a} = \bar{a}(\xi_1, \xi_2) \quad \underline{\text{KNOWN}}$$

We want a solution $S = S(q, t, \xi_1, \xi_2)$

satisfying:

- $S \geq 0$
- $|\partial^\alpha S| \leq C\left[|q| + |\xi_1| + |\xi_2| + 1\right]^m$ for $|\alpha| \leq 3$

(any m will do)

- $u = -\frac{1}{2}\partial_q S$ satisfies

$|u| \leq C\left[|q| + 1\right]$.

That's the PDE.

ON TO CONTROL THEORY...

THE LQR PROBLEM (SIMPLE 1D VERSION)

$q(t) \in \mathbb{R}^1$ POSITION

$u(t) \in \mathbb{R}^1$ CONTROL

DYNAMICS

- $q(t) = q_0$ AT TIME $t = 0$.

- $dq(t) = (aq(t) + u(t))\,dt + dW(t)$

[$W(t)$ = Brownian motion]

a = "GAIN"

GOAL (ROUGHLY SPEAKING) IS TO KEEP

$$\text{COST} = \int_0^T \{(q(t))^2 + (u(t))^2\}\, dt$$

AS SMALL AS POSSIBLE.

FLAVORS OF LQR:

- KNOWN a ← CLASSICAL

- UNKNOWN a, BUT GIVEN A BAYESIAN ← BAYESIAN PRIOR BELIEF, SPECIFIED BY A PROBABILITY DISTRIBUTION ON $[-a_{MAX}, +a_{MAX}]$.

- WE ARE TOLD THAT $a \in [-a_{MAX}, a_{MAX}]$
  BUT WE AREN'T GIVEN A PRIOR BELIEF

- a MIGHT BE ANY REAL NO.

↖ AGNOSTIC

FOR THE CLASSICAL & ~~AGS~~ BAYESIAN PROB'S

WE WORK IN A KNOWN PROBABILITY SPACE,

SO $E[\ldots]$ = EXPECTED VALUE

IS WELL-DEFINED.

Our goal is to pick a control strategy

to minimize

$E[\text{COST}]$.

For Agnostic Control Theory,

we Don't know which Probability Space

we are working in,

Because we aren't given a

Probability Distribution

for the Unknown a.

How to Formulate the Problem?

THE IDEA OF REGRET :

WE COMPARE THE PERFORMANCE OF

OUR STRATEGY σ

WITH THE PERFORMANCE OF AN

OPPONENT WHO KNOWS THE TRUE VALUE

OF a AND PLAYS PERFECTLY

(by applying CLASSICAL CONTROL THEORY).

If the TRUE VALUE OF THE GAIN IS a,

then the expected cost of our strategy σ

will be denoted by $EC(\sigma, a)$,

while the expected cost to our opponent

will be denoted by $EC(\text{Opponent}, a)$.

Additive Regret:

$$AR(\sigma, a) = EC(\sigma, a) - EC(\text{Opponent}, a)$$

May try to pick σ to minimize

$$\text{Max}\{ AR(\sigma, a) : a \in [-q_{MAX}, +q_{MAX}] \} \equiv AR_*(\sigma)$$

MULTIPLICATIVE REGRET

(AKA COMPETITIVE RATIO) :

$$MR(\sigma, a) = EC(\sigma, a) / EC(\text{Opponent}, a)$$

MAY TRY TO PICK σ TO MINIMIZE

$$\max\{ MR(\sigma, a) : a \in [-a_{max}, a_{max}] \} \equiv MR_*(\sigma)$$

HYBRID REGRET :

FIX AN ENTRANCE FEE $\delta > 0$.

$$HR(\sigma, a) = \frac{EC(\sigma, a) + \delta}{EC(\text{Opponent}, a) + \delta}$$

MAY TRY TO PICK σ TO MINIMIZE

$$\max\{ HR(\sigma, a) : a \in [-a_{MAX}, a_{MAX}] \} \equiv HR_*(\sigma)$$

(DEPENDS ON δ).

AGNOSTIC CONTROL THEORY PROBLEMS:

Find a strategy with least possible additive / multiplicative / hybrid regret.

TEXTBOOK CONTROL THEORY.

START THE GAME AT TIME t, POSITION q.

"COST TO GO" $S(q,t)$ IS MIN OF

$$\int_t^T \{(q(s))^2 + (u(s))^2\}\, ds$$

OVER ALL POSSIBLE CONTROL STRATEGIES

STARTING AT TIME t.

$$S(q,T) = 0$$

$$S(q,t) = \{q^2 + u^2\}\,dt + E\int_{t+dt}^{T} \{(q(s))^2 + (u(s))^2\}\,ds$$

$$= \{q^2 + u^2\}\,dt + E[S(q+dq,\ t+dt)].$$

Expand $S(q+dq,\ t+dt)$ to $\begin{bmatrix} 2^{\underline{nd}} \text{ order in } dq \\ 1^{\underline{rst}} \text{ order in } dt \end{bmatrix}$

Recall $dq = (aq+u)\,dt + dW$, so $E(dq) = (aq+u)\,dt$

$E(dq^2) = dt$.

So

$$S = (q^2 + u^2)\,dt + [S + \partial_t S\,dt + \partial_q S \cdot (aq + u)\,dt + \tfrac{1}{2}\partial_q^2 S] + o(dt)$$

with u picked to MINIMIZE RHS,

i.e.

$$\boxed{u = -\tfrac{1}{2}\partial_q S}$$

So

$$\boxed{0 = q^2 - \tfrac{1}{4}(\partial_q S)^2 + \partial_t S + aq\,\partial_q S + \tfrac{1}{2}\partial_q^2 S}$$

$$S = 0 \quad \text{for } t = T$$

The BELLMAN EQUATION

THE BELLMAN EQ. CAN BE EASILY SOLVED IN CLOSED FORM.

OPTIMAL CONTROL IS $u = -K(a,t)q$

$$\left[\begin{array}{l} K(a,t) \text{ SOLVES AN ODE,} \\ |K(a,t)| \leq C \text{ if } |a| \text{ bdd} \\ |K(a,t)| \leq Ca \text{ if } a \gg 1 \\ K(a,t) \to 0 \text{ as } a \to -\infty \end{array}\right]$$

So much for Classical Control Theory.

On To Bayesian Control

SUPPOSE THAT INITIALLY, THE UNKNOWN a HAS A KNOWN PROBABILITY DISTRIBUTION

$dProb(a)$ ON $[-a_{MAX}, +a_{MAX}]$.

Suppose we observe history $q(s), u(s)$ for times $s \in [0, t]$.

Our observations will change our beliefs about a.

WHAT'S OUR POSTERIOR PROB. DISTRIBUTION RE THE UNKNOWN a?

ANSWER:

$$dPROB^{(a)}_{POSTERIOR} = \phi(a) \cdot dPROB(a) \Big/ \text{NORMALIZING CONST.}$$

where $\phi(a)$ MAY BE COMPUTED FROM TWO NUMBERS ξ_1, ξ_2 DETERMINED FROM HISTORY UP TO TIME t

AS FOLLOWS ....

$$\xi_1 = \int_0^t (dq - u\,dt)$$

$$\xi_2 = \int_0^t q^2\,dt$$

KEY OBSERVATION :

WE NEEDN'T LOOK AT THE FULL HISTORY OF THE WORLD UP TO TIME t.

WE JUST NEED TO KNOW q, t, ξ_1, ξ_2.

COST TO GO IS NOW $S(q, t, \xi_1, \xi_2)$,

DEFINED AS THE LEAST POSSIBLE EXPECTED VALUE OF

$$\int_t^T \{ (q(s))^2 + (u(s))^2 \} ds,$$

GIVEN HISTORY UP TO TIME t, ASSUMING THAT AT TIME t

WE HAVE $q(t) = q$, $\xi_1(t) = \xi_1$, $\xi_2(t) = \xi_2$.

(OPTIMIZE OVER ALL CONTROL STRATEGIES u)

AGAIN WE GET A BELLMAN EQ. FOR S,

BUT INSTEAD OF AN ELEMENTARY PDE

WITH A CLOSED-FORM SOLUTION, WE

GET THE PDE STATED AT THE BEGINNING

OF THIS TALK.

AS IN THE SIMPLE CASE OF KNOWN a, WE HAVE

$$u = -\tfrac{1}{2}\partial_q S .$$

THM ON BAYESIAN CONTROL : ⟨ Given Prob, prior belief ⟩

The strategy given by $u = -\frac{1}{2}\partial_q S(q, t, \xi_1, \xi_2)$,
with S solving the Bellman eq as described above,
has the least possible expected cost, among all
possible strategies.

Moreover, any other strategy that does almost
as well makes almost the same decisions as the
$u = -\frac{1}{2}\partial_q S$ strategy.

So much for Bayesian control.

On to Agnostic control...

AGNOSTIC CONTROL I :

We are guaranteed that a belongs to $[-a_{MAX}, a_{MAX}]$, but we are NOT given a prior probability distribution for a.

Recall that we are trying to MINIMIZE WORST CASE REGRET (additive / multiplicative / hybrid)

THM ON AGNOSTIC CONTROL:

[FIX A NOTION OF REGRET (ADDITIVE/MULT./HYBRID).
FIX AN INTERVAL $[-q_{MAX}, q_{MAX}]$, & SUPPOSE
a MUST LIE IN THAT INTERVAL.]

Then there exists a prior prob. distribution $d\text{Prior}(a)$, CONCENTRATED ON FINITELY MANY PTS, such that the OPTIMAL BAYESIAN STRATEGY σ for $d\text{Prior}$ has the following properties.

(a) REGRET (σ, a) IS CONSTANT ON Support (dPrior)

(b) REGRET (σ, a) IS MAXIMIZED ON support (dPrior)
as a fn. of $a \in [-a_{MAX}, +a_{MAX}]$

(c) If σ' is any other strategy, then

$$\max_{a \in [-a_{MAX}, +a_{MAX}]} \text{REGRET}(\sigma, a) \leq \max_{a \in [-a_{MAX}, +a_{MAX}]} \text{REGRET}(\sigma', a)$$

A FEW REMARKS ON THE PROOF

START by assuming that the UNKNOWN a IS REQUIRED TO BELONG TO A FINITE SET A.

Under that assumption, we prove two key results:

① ANY STRATEGY WITH CLOSE-TO-OPTIMAL REGRET
IS CLOSE TO THE OPTIMAL BAYESIAN STRATEGY
FOR SOME PRIOR PROB. DISTRIBUTION SUPPORTED ON A.
(because two relevant convex sets are disjoint, hence
separated by a hyperplane)

---

② For the prior prob distribution in ①, the optimal
Bayesian strategy σ satisfies (a), (b), (c)
in statement of our Thm on AGNOSTIC CONTROL
(proved by induction on the number of elements of A)

NEXT, PASS FROM FINITE A

TO $[-a_{MAX}, a_{MAX}]$

by SIMPLY TAKING A TO BE

A FINE NET and PASSING TO THE LIMIT.

This gives all desired RESULTS EXCEPT

that we don't yet know that the relevant

prior prob. distribution on a has FINITE SUPPORT.

To obtain the finite support,

Recall that Property (a) [known by now]

asserts that

| Regret (σ, a) is Constant on Support of the Prior.

We show that

$\mathbb{R} \ni a \longmapsto$ Regret (σ, a) is Real-Analytic.

Hence, unless Regret (σ, a) is <u>constant</u>,

~~the counterexample~~ the support of the prior π a

<u>finite</u> subset of $[-a_{max}, +a_{max}]$

However, $a \longmapsto \text{REGRET}(\sigma, a)$ cannot be constant because one checks easily that

$$\text{REGRET}(\sigma, a) \longrightarrow \infty \quad \text{as} \quad a \to \infty.$$

Thus, our prior probability distribution is supported in a FINITE subset of $[-a_{MAX}, +a_{MAX}]$.

END OF STORY

AGNOSTIC CONTROL II :

WE KNOW ABSOLUTELY NOTHING

ABOUT a — IT COULD BE ANY REAL NUMBER.

GIVEN $\varepsilon > 0$, WE PRODUCE $a_{MAX} = a_{MAX}(\varepsilon)$,

TOGETHER WITH A PROCEDURE TO CONVERT AN

OPTIMAL STRATEGY FOR $a \in [-a_{MAX}, +a_{MAX}]$

INTO AN ε-ALMOST OPTIMAL STRATEGY FOR $a \in \mathbb{R}$.

THE CHALLENGE IS THAT IF $a >> 1$ IS REALLY HUGE,

THEN q WILL GROW TOO BIG, TOO FAST,

UNLESS WE RESPOND ALMOST IMMEDIATELY.

THE PROCEDURE :

① EARLY WARNING SYSTEM :

Make a quick decision re whether or not $a > a_{MAX}$.

② If we think $a \leq a_{MAX}$, then UNTIL FURTHER NOTICE, ~~execute~~ execute the optimal strategy for $a \in [-a_{MAX}, +a_{MAX}]$.

If instead we think $a > a_{MAX}$, then TAKE EXTREME MEASURES.

③ If $|q(t)|$ ever reaches a large threshold value Q, then conclude that a_{MAX} is very large positive, and TAKE EVEN MORE EXTREME MEASURES.

HOPE FOR THE FUTURE:

(1) The PDE PROBLEM WILL BE SOLVED

(2) Higher-dimensional and other variants of our simple agnostic control problems will be explored & solved. NEW ISSUES ARISE.

(3) Practical applications someday ?

THANK YOU FOR LISTENING!