An Algebraic Approach To Optimization

Nigel Boston

Departments of ECE and Math

University of Wisconsin

A Math Problem

- For given $a,b \in \mathbf{R}$, but varying $c,d \in \mathbf{R}$, among all polynomials of the form $f(z) = 1+az^{-1}+bz^{-2}+cz^{-3}+dz^{-4}$ which one has the smallest largest root?

- Polynomial $f(z)$ is called the pole radius minimizer of degree 4 of $1+az^{-1}+bz^{-2}$.

Solution to the Math Problem

(c,d) is one of the following:

(1) $(ab/2 - a^3/8,\ (a^2-4b)^2/64)$

(2) $((-a^3 \pm a^2(a^2-8b)^{1/2} + 4ab)/8,\ (-a^4 \pm a^3(a^2-8b)^{1/2} + 4a^2b + 8b^2)/32)$

(3) $((-9a^3 \pm 3^{1/2}(3a^2-8b)^{3/2} + 36ab)/72,\ (-27a^4 \pm 3^{3/2}a(3a^2-8b)^{3/2} + 108a^2b - 72b^2)/864)$

The ± are chosen to be the same in c and d.

Simple Look-Up

This leaves finitely many polynomials to check but it's actually even simpler.

If $8b > 3a^2$, then (1) is best.

If not and $b>0$, then (2) is best, using + if $a>0$ and - if $a<0$.

If not and $b<0$, then (3) is best, using + if $a>0$ and - if $a<0$.

Where Does This Come From?

- There's a very simple idea behind this.
- The pole radius minimizers have roots of the form $(\alpha,\alpha,\beta,\beta)$, $(\alpha,\alpha,-\alpha,\beta)$, $(\alpha,\alpha,\alpha,\beta)$.
- In each case, write a,b,c,d in terms of α,β and then solve for c,d in terms of a,b.

Pipelined IIR Filter Architecture

From a question of Naresh Shanbhag (UIUC).

The transfer function of an IIR filter is B(z)/A(z) (A,B polynomials in z^{-1}).

If we can find a "stable" D(z) such that the first few coefficients of A(z)D(z) are sums of very few powers of 2, then the filter is pipelined and throughput improved.

Mathematical Translation

Suppose we specify that

$A(z)D(z) = 1 + c_1 z^{-1} + \ldots + c_M z^{-M} + \ldots$

($c_1,\ldots,c_M$ in some small set of powers of 2, etc.)

Since A(z) is given, we solve to get

$D(z) = 1 + d_1 z^{-1} + \ldots + d_M z^{-M} + \ldots$

For stability, we want D(z) to be a pole radius minimizer whose coefficients start that way.

Practical Example

6th order low-pass Butterworth filter with 0.3 cutoff frequency. MATLAB gives:

$A(z) = 1 - 2.3797z^{-1} + 2.9104z^{-2} - 2.0551z^{-3} + 0.8779z^{-4} - 0.2099z^{-5} + 0.0218z^{-6}$

With $c_1=-2$, $c_2=2$, $c_3=-1$, $c_4=0.25$, $c_5=0$, $c_6=0$, we obtain a stable D(z) with largest pole radius 0.6894.

Comparison with Other Methods

MAC counts multiplier-adder pairs, A adders.

Scattered lookahead: 18 MAC (0.8085).

Minimum order augmentation: 18 MAC (0.7807).

Sum of powers of 2 (SPOT): 14 MAC + 11 A (0.7177).

Our method: 6 MAC + 5 A (0.6894).

Belgian Chocolate Problem

A problem in controller design for which Blondel offered 1 kg of chocolate.

Given a process parameter δ, find stable polynomials (all roots in the left half-plane) $x(s)$, $y(s)$, $z(s)=(s^2-2\delta s+1)x(s) + (s^2-1)y(s)$ with $\deg y < \deg x$.

Blondel's Problem

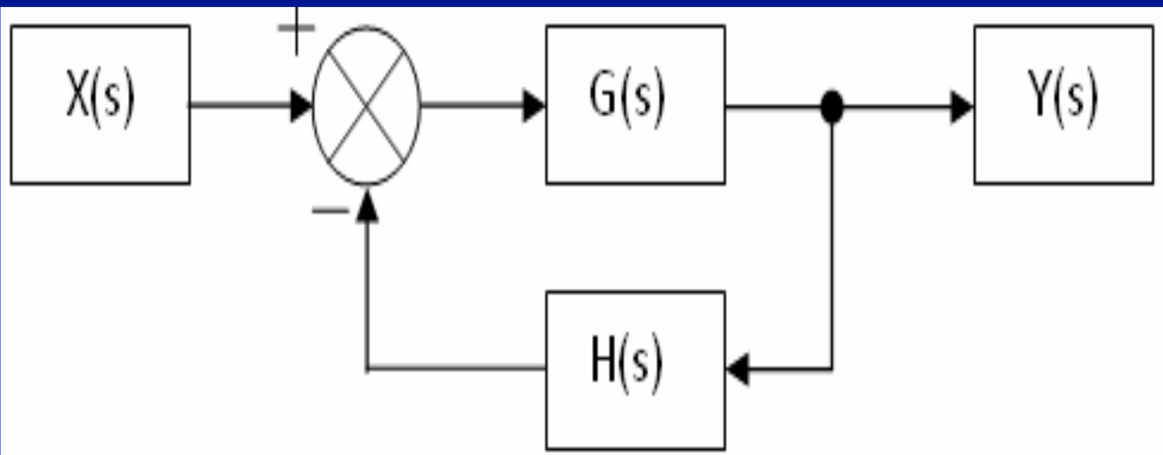
Consider plant $P(s) = (s^2-1)/(s^2-2\delta s+1)$

This is open-loop unstable

Seek stabilization by a stable minimum phase controller $C(s) = y(s)/x(s)$ (deg $x >$ deg y)

Putting $C(s)$ and $P(s)$ in closed-loop series, we compute transfer function $= PC/(1+PC)$, whose denominator is $z(s)$

Closed-Loop Transfer Function



The summing node and the *G*(*s*) and *H*(*s*) blocks can all be combined into one block, which would have the following transfer function:

$$\frac{Y(s)}{X(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

Applicable?

Control problems with near cancellation of unstable poles and zeros (just as in the Belgian chocolate problem) arise in physically relevant engineering problems, e.g. the X-29 prototype aircraft design problem or Klein's bicycle design problem.

Ours is the simplest nontrivial stabilization you can cook up.

Work So Far

- 1994 - Blondel offers 1kg of chocolate for a solution for $\delta = 0.9$ and 1kg for full solution
- 2002 - Patel et al gave solution for $\delta = 0.9$ with x,y of degree 11
- 2005 - Burke et al gave solution for $\delta = 0.9$ with deg $x = 3$, y constant
- 2007 -Chang et al gave soln for δ=0.973974

Example for Degree 6

Say deg x = 6. Chang-Sahinidis had the best solution: $\delta = 0.96292177890276$, $x(s) =$

$s^6 + 1.9267063716832s^5 + 2.7125040416507s^4 + 3.2971535543909s^3 + 2.4444879197567s^2 + 1.4102519904753s + 0.7321239653705$,

$y(s) = 1.1928111395529s^2 + 0.0002957682513s + 0.7321239653705$.

New Algebraic Method

They used global optimization algorithms to obtain ever better solutions.

The algebraic approach is instead to identify what these solutions are converging to. This end-run gives the optimal δ and solutions as close as we like to the optimal.

Optimal Solution for n=6

$x(s) = (s^2+2\delta\, s+1)\,(s^2 + 2\delta^2 - 1)^2 = s^6 + 1.92594791192221s^5 + 2.70927535943690s^4 + 3.29197530940740s^3 + 2.43968092303150s^2 + 1.40672307006120s + 0.730405563594600,$

$y(s) = (12\delta^4 - 12\delta^2 + 2)s^2 + (4\delta^4 - 4\delta^2 + 1) = 1.19121669078370s^2 + 0.730405563594600,$

$z(s) = s^8$, where δ is the largest real root of $16x^6 - 16x^4 + 1 = 0$, $\approx 0.962973955961027$

Arbitrarily Good Solutions

This yields stable polynomials x,y,z for δ as close to 0.962973955961027 as we like

Simply take the optimal solution and deform the coefficients ever so slightly but in the right direction (increase them) and by carefully chosen amounts

Computing Optimal Solutions

The optimal solutions have most of their roots on the imaginary axis with many repeated

So x,y,z have many factors of the form s^2+k

For n=6, try $x(s) = (s^2+2\delta\, s+1)(s^2+a)^2$,

$y(s) = bs^2 + c$, $z(s) = s^8$

and obtain a system of polynomial equations in a,b,c,δ to solve

Details

$z(s) = (s^2 - 2\delta\, s + 1)x(s) + (s^2 - 1)y(s) \;\Rightarrow\; s^8 =$

$(s^2 - 2\delta\, s + 1)(s^2 + 2\delta\, s + 1)(s^2 + a)^2 + (s^2 - 1)(bs^2 + c)$

$= ((s^2 + 1)^2 - 4\delta^2\, s^2)(s^2 + a)^2 + (s^2 - 1)(bs^2 + c)$

Equating coefficients of s^6, s^4, s^2, s^0 gives:

$2 - 4\delta^2 + 2a = 0,$

$1 + 2a(2 - 4\delta^2) + a^2 + b = 0,$

$2a + a^2(2 - 4\delta^2) + c - b = 0, \quad -c + a^2 = 0.$

4 Equations in 4 Unknowns

These are easily solved:

$a = 2\delta^2 - 1$,

$c = a^2 = (2\delta^2 - 1)^2$,

$b = -(1 + 2a(2 - 4\delta^2) + a^2) = 12\delta^4 - 12\delta^2 + 2$.

Plugging back into the remaining equation gives $16\delta^6 - 16\delta^4 + 1 = 0$

Belgian Chocolate?

Not yet.

For a given degree 2k, the method yields k+1 polynomial equations in k+1 variables (δ and certain coefficients of x(s), y(s), z(s)). These can be solved exactly for small k and numerically for larger k.

I announce new world record $\delta = 0.976462$.

Soln with $\delta = 0.97646152004586$

$x(s) = (s^2+2\delta\, s+1)(s^2+a)^3(s^2+b)^2$

$y(s) = \mu(s^2+c)^2(s^2+d)$

$z(s) = s^{12}(s^2+e)$

$z(s)=(s^2-2\delta\, s+1)x(s) + (s^2-1)y(s)$ yields

7 polynomial equations in a,b,c,d,e,μ,δ

Optimal Solution for n=12

$a = 1.199452572585249,$

$b = 0.3120456711810086,$

$c = 0.4390239521999522,$

$d = 0.2824068421308690,$

$e = 2.408540659596713,$

$\mu = 3.086980745129855,$

$\delta = 0.9764615200458$6

abc Theorem for Polynomials

Suppose that a(s) + b(s) = c(s).

max(deg a,deg b, deg c)<#distinct roots of abc

Proven by Stothers and Mason, it implies Fermat's Last Theorem for polynomials, and is the motivation for the fundamental abc conjecture of number theory

abc Near-Misses

Looking at the optimal solution for n=12, it is a+b = c, with a,b,c of degree ≤ 7 in s^2

The distinct roots of abc are 0,1,-a,-b,-c,-d,-e, and the two roots of $(x+1)^2-4\delta^2 x=0$

In fact, in all optimal solutions found so far, the maximal degree is 2 less than the number of distinct roots

Deeper Interpretation

Examples where the maximal degree is 1 less than the number of distinct roots are related to Grothendieck's dessins d'enfants, which describe covers of P^1 ramified at 3 points

Where the difference is 2, rather than 1, is related to covers of P^1 ramified at 4 points, described by Hurwitz schemes

Further Plans

To use this interpretation to suggest higher degree identities, seeking a sequence of these with ever larger δ

To obtain a systematic theory of relaxation from the optimal solution

Conclusions

The Belgian chocolate problem is handled by an "algebraic end-run", obtaining the limiting solution rather than creeping towards it by multiple numerical algorithms

Polynomials with repeated roots solve some optimization problems ->algebraic solutions

Limiting solutions of the Belgian chocolate problem <-> covers of P^1 ramified at 4 pts

Questions?