Smoothing Splines with Derivative Constraints

Rochana Meegaskumbura

PhD Final Examination
Department of Mathematics & Statistics
Texas Tech University
Lubbock
Splines

- A spline is a special function defined piecewise by polynomials of same degree, with continuity constraints on derivatives and possibly constraints on end points.

Given a set of data points \((x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)\), we can have two types of splines:

- **Interpolating Splines:**
  Spline function \(S(x)\) satisfies \(y_i = S(x_i)\) for \(i = 1, 2, \ldots, n\).

- **Smoothing Splines:**
  Smoothing data using a spline function \(S(x)\). Deviation of the spline function \(S(x)\) from the data points should minimize a cost function.
Control theoretic splines are smoothing splines with constraints written in terms of a linear differential equation and the smoothing curve between data points is obtained by an input to a given linear system.
Previous Work

1. Theory on optimal control and monotonic smoothing splines:
   Control Theoretic Splines: Optimal Control, Statistics, and Path Planning
   Magnus Egerstedt and Clyde Martin

2. A quadratic programming approach is developed in
   Quadratic Programming for Monotone Control Theoretic Splines
   Masaaki Nagahara, Clyde F. Martin and Yutaka Yamamoto
Problem Definition

Given a linear time invariant system

\[ \dot{x} = Ax + Bu, \quad y = Cx, \quad x(0) = x_0 \]

where,

\[ A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \]

with constraints

\[ y(0) = y(1) = 0, \quad y(t) \geq 0 \quad \text{and} \quad y''(t) < 0. \]

Find the control and initial condition pair \((u(t), x_0)\) that minimizes the following cost function

\[ J(u, x_0) = w_1 x_0^T x_0 + w_2 \int_0^1 u(t)^2 \, dt + \sum_{i=1}^{N} \omega_i (y(t_i) - \alpha_i)^2, \]

for a given data set \((t_i, \alpha_i), \ i = 1, \ldots, N\), and weights \(\omega_1, \ldots, \omega_N, w_1, w_2\).
Lemma

Prove $H = \left\{ (u(t), x_0) : y(t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-s)}Bu(s) \, ds, \right.$

$\left. y(0) = y(1) = 0, y(t) \geq 0, y''(t) < 0 \right\} \neq \emptyset$.

Let $u(t) = -2t$ and $x_0 = (0, 1)$.

Then $y(t) = t(1 - t^2)$ and we have

$y(0) = y(1) = 0, y(t) \geq 0, y''(t) < 0$.

Therefore, $H \neq \emptyset$. 
The set $H$ is a convex, nonempty and closed cone

**To show convexity:**

Let $(u_i, x_0^i) \in H$ where $i = 1, 2$.

For $\alpha, \beta \geq 0$, we have the pair $p = (\alpha u_1 + \beta u_2, \alpha x_0^1 + \beta x_0^2)$ satisfying the constraints.

Therefore, $p \in H$.

Hence $H$ is a convex cone.
The set $H$ is a convex, nonempty and closed cone

To show it is closed:

Consider a collection of control and initial condition data pairs 
\[ \{(u_i, x^i_0)\}_{i=1}^{\infty} \subset H. \]

Where each individual pair satisfies $x^i_0 \to \hat{x}$ and $u_i \to \hat{u}$ as $i \to \infty$.

Then, 
\[
\lim_{i \to \infty} y_i(t) = \lim_{i \to \infty} Ce^{At}x^i_0 + \int_{0}^{t} Ce^{A(t-s)}B \lim_{i \to \infty} u_i(s) \, ds.
\]

Therefore each individual pair will satisfy the constraints.

We can pass the limit through the integral due to the compactness of $[0, 1]$ and continuity.

Then we have the pair $(\hat{u}, \hat{x})$ satisfying the constraints we have the limit $(\hat{u}, \hat{x}) \in H$.

Therefore it is closed.

We have already proven that it is non empty.
Another Useful Theorem

\[ J(u, x_0) = w_1 x_0^T x_0 + w_2 \int_0^1 u(t)^2 \, dt + \sum_{i=1}^N \omega_i (y(t_i) - \alpha_i)^2, \]
the cost function is convex in both \( x_0 \) and \( u \).

Hence has a unique minimum \( \in H \)

The optimal control yields a cubic spline in the output of the control system. [1]
Finding Optimal Control \((u^*, x_0^*)\)

Consider the cost function with weights \(w_1, w_2\) and \(\omega_i\)'s.

\[
J(u, x_0) = w_1 x_0^T x_0 + w_2 \int_0^1 u(t)^2 \, dt + \sum_{i=1}^{N} \omega_i \left( y(t_i) - \alpha_i \right)^2.
\]

- A simple control of the form \(u(t) = -at + b\) was tested with initial condition \(x_0 = (0, c)\).
- The coefficients \(a, b\) and \(c\) were found to be very sensitive to the end point constraint \(y(1) = 0\).
- An algorithm based on the hill climbing optimization was used.
- Cost function with \(w_1 = 0, w_2 = 10^{-5}\) [1] was used.
Finding Optimal Control \((u^*, x_0^*)\): The Algorithm

initialize \(a = 2, b = 0, c = 1\)
initialize \(count = 0, iter = 0\)
define control: \(u_t = -at + b\)
define \(x = (0, c)\)
define accuracy: \(\epsilon = 10^{-5}\)
define iteration limits:
\(L_1 = 1000\)
\(L_2 = 1000\)
define Local min. of cost: \(J_{lm}\)
define Global min. of cost: \(J_{gm}\)
define Current value of cost: \(J\)
define Step size: \(\Delta\)

while \(J_{gm} > J\) or \(count = 0\) do
  while \(J_{lm} > J\) or \(iter = 0\) do
    set \(J_{lm} = J\)
    Find the search dirn of \(a\).
    move \(a\) \((a = a \pm \Delta)\)
    change \(b\) until constraints are satisfied

  evaluate cost \(J\)
  iter = iter + 1
  if \(iter > L_1\) then
    break
  end if
end while
Set \(J_{lm} = J_{gm}\)
Find the search dirn of \(c\).
move \(c\) \((c = c \pm \Delta)\)
change \(b\) until constraints are satisfied
evaluate \(J\)
count = count + 1
if \(J_{gm} \leq \epsilon\) or \(count > L_2\) then
  break
end if
Finding Optimal Control \((u^*, x_0^*)\)

- Algorithm was first tested with a sine function: \(\sin(\pi t)\) where \(t \in [0, 1]\)
- Optimal values obtained from the hill climbing algorithm
  - Control \(u = -at + b\)
  - Optimal parameters \(a = 2.09, b = -4.59\) and \(x = (0, 3.33)\)
  - Minimum cost \(J_{min} = 0.1995\)
  - \(y(t) = 3.33t + 0.5(2.09t + 4.59)t^2 + (-2.09t + 4.59)t^2\)
Approximating with a Single Control

Figure: Approximating with the curve produced by a single control
Shifting the curve to the left is found impossible by coefficient tuning of the tested control with the optimization algorithms.

Therefore a piecewise control was implemented.
Define control

\[
u(t) = \begin{cases} 
-a_i(t - t_i) + b_i, & t_i \leq t < t_{i+1} \\
0, & \text{otherwise}
\end{cases}
\]

- \( b_1 = 0, \ b_i = u_{i-1}(t_i) \) for \( i = 2, \ldots, 10 \),
- \( a_i \) is the gradient of each curve \( i = 1, 2, \ldots, 10 \),
- \( t = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1] \)
- Initially all gradients were set to 2.
- Cost function with \( w_1 = 0, \ w_2 = 10^{-5} \) [1] was used.
Finding Optimal Control \((u^*, x_0^*)\): The Algorithm

**initialize** \(a = [2, 2, 2, \ldots, 2]\), \(b_1 = 0\), \(c = 1\)

**initialize** count = 0, iter = 0

**define**: \(t = [0.1, 0.2, 0.3, \ldots 1]\)

**define** control:

\[ u(t) = -a_i(t - t_i) + b_i, \quad t_i \leq t < t_{i+1} \]

**define** \(b_i = u_{i-1}t_i\) for \(i > 1\)

**define** \(x = (0, c)\)

**define** accuracy: \(\epsilon = 10^{-5}\)

**define** limits: \(L_1 = 1000\), \(L_2 = 1000\)

**define** Local min. of cost: \(J_{lm}\)

**define** Global min. of cost: \(J_{gm}\)

**define** Current cost: \(J\)

**define** Step size: \(\Delta\)

Find \(a(10)\) to satisfy the constraint \((y(1) = 0)\)

while \(J_{gm} > J\) or count = 0 do
  for \(i=1:9\) do
    while \(J_{lm} > J\) or iter = 0 do
      set \(J_{lm} = J\)
      Find the search dirn of \(a(i)\).
      move \(a(i)\)
      change \(a(10)\) until constraints are satisfied
    end while
  end for

evaluate cost \(J\)

iter = iter + 1

if \(iter > L_1\) then
  break
end if

Check global convexity

\(y(t) < 0\) for all \(t\)

if \(i > 2\) and \(|a(i) - a(i-1)| > 1\) then
  break
end if

end while

end while

end for

Set \(J_{lm} = J_{gm}\)

Find the search dirn of \(c\).

move \(c\) \((c = c \pm \Delta)\)

change \(a(10)\) until constraints are satisfied

evaluate \(J\)

\(count = J + 1\)

if \(J_{gm} \leq \epsilon\) or \(count > L_2\) then
  break
end if

end while
Figure: Curve produced by the piecewise control approximating smooth sine curve
The Optimal Piecewise Output

\[ y(t) = ct - 0.5a_1 t^3 - 0.5(a_3 t - 2.80)t^2 - 0.5(a_4 t - 1.60)t^2 \\
     - 0.5(a_5 t - 0.640)t^2 - 0.5(a_6 t - 0.260)t^2 - 0.5(a_7 t - 1.34)t^2 \\
     - 0.5(a_8 t - 2.46)t^2 - 0.5(a_9 t - 3.74)t^2 - 0.5(a_{10} t - 5.00)t^2 \]

Where \[ a_i = \begin{cases} 
    a_i, & t_i \leq t < t_{i+1} \\
    0, & \text{otherwise} \end{cases} \]

\[ c = 3, \quad a = [1.6, 1.6, 15.6, 11.6, 9.2, 7.4, 5.6, 4.0, 2.4, 1.0] \]

Minimum cost is 0.0042
Random noise introduced in to a sine function
- Mean = 0
- Standard deviation = 0.1
Figure: Approximating data with added noise with the curve produced by a single control.
The Optimal Output

- $u = -at + b$

Optimal values

- $a = -2.4$
- $b = -4.26$
- $x = (0, 3.32)$
- Minimum cost = 1.4505
- $y(t) = 3.32t - 0.5(2.40t + 4.26)t^2$
Comparison of output

- control: initial
- control: optimal
- sine with noise

Output $y(t)$ vs. Time $t$
The Optimal Piecewise Output

\[ y(t) = ct - 0.5a_1 t^3 - 0.5(a_3 t - 1.13)t^2 - 0.5(a_4 t - 2.32)t^2 \]
\[ - 0.5(a_5 t - 0.6)t^2 - 0.5(a_6 t - 2.9)t^2 - 0.5(a_7 t - 5.32)t^2 \]
\[ - 0.5(a_8 t - 4.34)t^2 - 0.5(a_9 t - 6.86)t^2 - 0.5(a_{10} t - 5.60)t^2 \]

Where \[ a_i = \begin{cases} a_i, & t_i \leq t < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \]

\[ c = 3, \quad a = [1.0, 12.3, 12.9, 1.0, 8.3, 12.9, -0.8, 13.0, -1, 0.4] \]

Minimum cost is 1.4684
Sine Function with Added Noise: Piecewise Control with global convexity adjustment

Comparison of output $y(t)$

- control: initial
- control: optimal
- sine with noise
- sine curve
The Optimal Piecewise Output

\[ y(t) = ct - 0.5a_1 t^3 - 0.5(a_2 t + 4.55) t^2 - 0.5(a_3 t + 4.51) t^2 \]
\[ - 0.5(a_4 t + 4.69) t^2 - 0.5(a_5 t + 4.45) t^2 - 0.5(a_6 t + 4.40) t^2 \]
\[ - 0.5(a_7 t + 4.34) t^2 - 0.5(a_8 t + 4.27) t^2 - 0.5(a_9 t + 4.19) t^2 \]
\[ - 0.5(a_{10} t + 5.0) \]

Where \( a_i = \begin{cases} 
    a_i, & t_i \leq t < t_{i+1} \\
    0, & \text{otherwise} 
\end{cases} \)

\( c = 3, \quad \mathbf{a} = [47.4, 1.9, 2.1, 1.5, 2.1, 2.2, 2.3, 2.4, 2.5, 1.6] \)

Minimum cost is 1.4768
Scattered Data

- Absolute value of a random data set was generated with:
- Mean = 0
- Standard deviation = 1
The Optimal Output

- \( u = -at + b \)

- Optimal parameters
  - \( a = -2.03 \)
  - \( b = -4.59 \)
  - \( x = (0, 3.30) \)
  - Minimum cost = 50.3110
  - \( y(t) = 3.30t - 0.5(2.03t + 4.59)t^2 \)
Scattered Data: Piecewise Control

Comparison of output $y(t)$

- control: initial
- control: optimal
- scattered data

Output $y(t)$ vs. Time $t$
The Optimal Piecewise Output

\[ y(t) = ct - 0.5a_1 t^3 - 1.5(a_2 t + 4.76) t^2 - 0.5(a_3 t + 4.64) t^2 \\
- 0.5(a_4 t + 4.85) t^2 - 0.5(a_5 t + 5.13) t^2 - 0.5(a_6 t + 5.23) t^2 \\
- 0.5(a_7 t + 5.29) t^2 + (a_8 t + 5.36) - 2(a_9 t + 5.36) \\
- 0.5(a_{10} + 3.20) \]

Where \( a_i = \begin{cases} a_i, & t_i \leq t < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \)

\( c = 3, \quad a = [49.8, 2.2, 2.8, 2.1, 1.4, 1.2, 1.1, 1.0, 1.0, 3.4] \)

Minimum cost is 49.5417
sequential quadratic programming (SQP) methods find an approximate solution of a sequence of quadratic programming (QP) subproblems.

quadratic model of the objective function is minimized subject to the linearized constraints.
Using MATLAB function "fmincon"

- This is a function in MATLAB to optimize non-linearly constrained optimization problems based on SQP methods.
- syntax used is
  \[ x = \text{fmincon}(\text{fun}, x0, A, b, Aeq, beq, lb, ub, nonlcon, options). \]
Smooth Data: Single control

Figure: Approximating curve for smooth data for the single control
The optimal values

- Optimal control \( u = -0.0026t - 7.7363 \)
- Optimal initial condition \([0, 3.8695]\)
- Approximating curve
  \[ y(t) = 3.87t + 0.5(0.265e^{-2t} + 7.74) + t^2(-0.265e^{-2t} - 7.74) \]
- Minimum cost \( J_{min} = 0.0721 \)
Approximating curve for smooth data for the piecewise control

![Graph showing approximating curve and data set: sine curve.](image)
The optimal control is given by

\[ u(t) = \begin{cases} 
-a_i(t - t_i) + b_i, & t_i \leq t < t_{i+1} \\
0, & \text{otherwise}
\end{cases} \]

where \( b_i = u_{i-1} t_i \) and

\[ a = [3.157, -22.17, -14.90, -9.7594, 1.8409, 1.08, 0.5752, 0.2216, -0.0359] \]

The approximating curve is

\[ y(t) = ct + .5(a_1 t + 11.9)t^2 + t^2(-a_1 t - 11.9) + .5(a_2 t + 14.4)t^2 + t^2(a_2 t - 14.4) + .5(a_3 t + 13.0)t^2 + t^2(a_3 t - 13.0) + .5(a_4 t + 11.4)t^2 + t^2(a_4 t - 11.4) + .5(a_5 t + 6.78)t^2 + t^2(-a_5 t - 6.78) + .5(a_6 t + 7.16)t^2 + t^2(-a_6 t - 7.16) + .5(a_7 t + 7.46)t^2 + t^2(-a_7 t - 7.46) + .5(a_8 t + 7.71)t^2 + t^2(-a_8 t - 7.71) + .5(a_9 t + 7.92)t^2 + t^2(-a_9 t - 7.92) + .5(a_{10} t + 8.09)t^2 + t^2(a_{10} t - 8.09) \]

Minimum cost \( J_{\text{min}} = 0.0966 \)
Approximating curve for noisy data for the single control

Output $y(t)$

Data set: sine with noise

Time $t$
Optimal values

- Optimal control $u = -1.1367t - 6.1572$
- Optimal initial condition $[0, 3.647]$
- The approximating curve is
  \[ y(t) = 3.65t + 0.5(1.14t + 6.16)t^2 + t^2(-1.14t - 6.16) \]
- Minimum cost $J_{\text{min}} = 1.3701$
Approximating curve for noisy data for the piecewise control

![Approximating curve for noisy data for the piecewise control](image)

- **Output y(t)**
- **Time t**

- **Approximating curve**
- **Data set: sine with noise**
The optimal control is given by

\[ u(t) = \begin{cases} 
-a_i(t - t_i) + b_i, & t_i \leq t < t_{i+1} \\
0, & \text{otherwise}
\end{cases} \]

where \( b_i = u_{i-1} t_i \) and

\[ a = [2.9481, -27.7586, -17.9584, -11.7658, -0.8918, -0.3705, -0.8424, \ldots] \]

The optimal initial condition is \([0, c]\) where \( c = 4.2220\)

\[ y(t) = ct + .5(a_1 t + 14.5) t^2 + t^2(-a_1 t - 14.5) + .5(-a_2 t + 17.6) t^2 + t^2(a_2 t - \\
+ .5(a_3 t + 15.6) t^2 + t^2(-a_3 t - 15.6) + .5(-a_4 t + 13.8) t^2 + t^2(a_4 t - \\
+ .5(a_5 t + 9.42) t^2 + t^2(-a_5 t - 9.42) + .5(-a_6 t + 9.16) t^2 + t^2(a_6 t - \\
+ .5(a_7 t + 9.44) t^2 + t^2(-a_7 t - 9.44) + .5(-a_8 t + 9.66) t^2 + t^2(a_8 t - \\
+ .5(a_9 t + 9.83) t^2 + t^2(-a_9 t - 9.83) + .5(-a_{10} t + 9.97) t^2 + t^2(a_{10} t - \ldots) \]

Rochana Meegaskumbura

Smoothing Splines with Derivative Constraints
Optimal Values

- Optimal control \( u = 4.2627t - 13.9465 \)
- Optimal initial condition \([0, 4.8419]\)
- The approximating curve is 
  \( y(t) = 4.84t + 0.5(-4.26t + 13.9)t^2 + t^2(4.26t - 13.9) \)
- The minimum cost \( J_{min} = 49.149 \)
Approximating curve for scattered data for the piecewise control
Optimal Values

\[ u(t) = \begin{cases} 
-a_i(t - t_i) + b_i, & t_i \leq t < t_{i+1} \\
0, & \text{otherwise}
\end{cases} \]

where \( b_i = u_{i-1}t_i \) and


The optimal initial condition is \([0, c]\) where \( c = 4.9211 \)

\[ y(t) = ct + 0.5(a_1 t + 14.5)t^2 + t^2(-a_1 t - 14.5) + 0.5(a_2 t + 14.7)t^2 + t^2(a_2 t + 14.7) + 0.5(-a_3 t + 16.3)t^2 + t^2(-a_3 t - 16.3) + 0.5(a_4 t + 17.0)t^2 + t^2(a_4 t + 17.0) + 0.5(a_5 t + 17.3)t^2 + t^2(a_5 t + 17.3) + 0.5(a_6 t + 17.4)t^2 + t^2(a_6 t + 17.4) + 0.5(a_7 t + 16.2)t^2 + t^2(a_7 t + 16.2) + 0.5(a_8 t + 15.2)t^2 + t^2(a_8 t + 15.2) + 0.5(a_9 t + 14.4)t^2 + t^2(a_9 t + 14.4) + 0.5(a_{10} t + 7.04)t^2 + t^2(a_{10} t + 7.04) \]

Minimum cost \( J_{\text{min}} = 45.5202 \)
Algorithm was also tested for arbitrary start and end points different than 0 and 1.

Figure: Here the starting point is 0.8 and the end point is 1.
**Figure:** Here the starting point is 0 and the end point is 0.3
**Figure:** Here the starting point is 0.3 and the end point is 0.8
The classical smoothing splines tend to overshoot at the corners when approximating this type of curves.

**Figure:** Approximating a curve with corners.
Conclusions

- It is proven that the feasible set for the problem is non empty.
- It is proven that the optimization problem has a unique minimum.
- Smooth data is better approximated with the piecewise curve as the input.
- Noisy and scattered data is approximated equally well with both the single and the piecewise control as input.
Thank you

- Dr. Clyde Martin
- Dr. Bijoy Ghosh
- Dr. Kevin Long
- Mr. Mervyn Ekanayake for helping me with the presentation
- All others who helped me in numerous ways