Piecewise-Linear Approximation: One Dimensional

Eduardo Camponogara

Department of Automation and Systems Engineering
Federal University of Santa Catarina

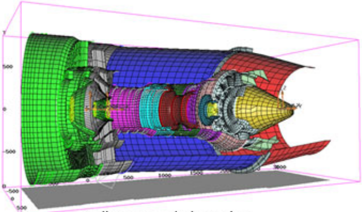
October 2016

Summary

Surrogate Model

Engineering problems involve experiments, and simulations, to evaluate objectives and constraints which are functions of several variables.

Optimal design of jet engine.

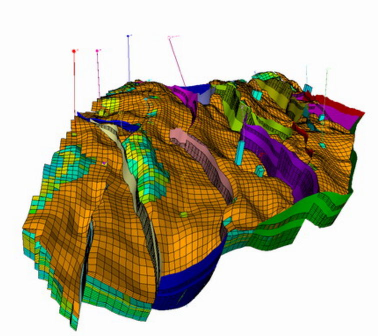


Rolls-Royce whole engine structural model (VIVACE project)

Surrogate Model

Engineering problems involve experiments, and simulations, to evaluate objectives and constraints which are functions of several variables.

Advanced recovery of oil and gas reservoirs.



Surrogate Model

Challenges:

- In real-world problems, a simulation run can take several minutes, hours and even days.
- Design optimization and case studies may become impractical due to the potential need of thousands of simulations.

Surrogate Model

Alternative:

- Propose approximate models ("Surrogate Models") that emulate the behavior of systems and simulators, however at a low computational cost.
- Surrogate Models are built from data, since the simulation model might not be known or is too complicated to be expressed in explicit form.
- Knowledge of the input-output behavior is known or given.

Surrogate Model

Alternative:

- Propose approximate models ("Surrogate Models") that emulate the behavior of systems and simulators, however at a low computational cost.
- Surrogate Models are built from data, since the simulation model might not be known or is too complicated to be expressed in explicit form.
- Knowledge of the input-output behavior is known or given.

Surrogate Model

Alternative:

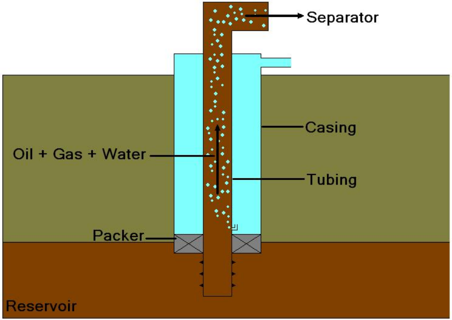
- Propose approximate models ("Surrogate Models") that emulate the behavior of systems and simulators, however at a low computational cost.
- Surrogate Models are built from data, since the simulation model might not be known or is too complicated to be expressed in explicit form.
- Knowledge of the input-output behavior is known or given.

Surrogate Model

"All models are wrong, the practical question is how wrong do they have to be to not be useful." George Box

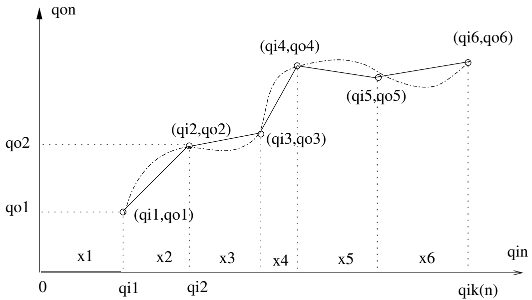
Piecewise-Linear Approximation

Models are obtained from linear (affine) combination of input-output data: $q_{oil}(q_{inj})$.



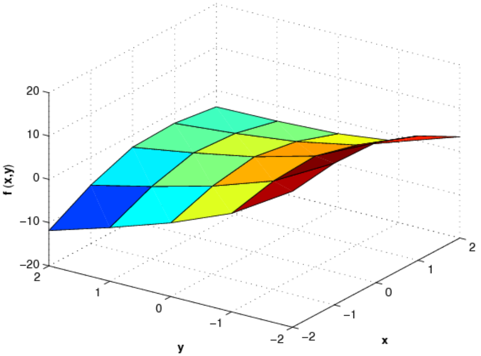
Piecewise-Linear Approximation

- Models are obtained from linear (affine) combination of input-output data: $q_{oil}(q_{inj})$.
- Data: $\{(q^1_{inj}, q^1_{oil}), (q^2_{inj}, q^2_{oil}), \ldots, (q^n_{inj}, q^n_{oil})\}$.



Piecewise-Linear Approximation

Function $f(x, y)$ with a two dimensional domain.



Piecewise-Linear Approximation

Question:

- How does one represent piecewise-linear functions in mathematical programming?

Several Models:

- CC (Convex Combination)
- Inc (Incremental)
- DCC (Disaggregated Convex Combination)
- Log (Logarithmic Convex Combination)
- DLog (Disaggregated Logarithmic Convex Combination)
- Multiple Choice
- SOS2 (Specially Ordered Sets of Type 2)

Piecewise-Linear Approximation

Question:

- How does one represent piecewise-linear functions in mathematical programming?

Several Models:

- CC (Convex Combination)
- Inc (Incremental)
- DCC (Disaggregated Convex Combination)
- Log (Logarithmic Convex Combination)
- DLog (Disaggregated Logarithmic Convex Combination)
- Multiple Choice
- SOS2 (Specially Ordered Sets of Type 2)

Summary

Motivation

Piecewise-Linear Models

Convex Combination (CC)

Data: $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$.

$$x = \sum_{i=0}^{n} \lambda_i x_i \qquad y = \sum_{i=0}^{n} \lambda_i y_i$$

$$1 = \sum_{i=0}^{n} \lambda_i \qquad \lambda_i \geq 0,\ i = 0, \ldots, n$$

$$1 = \sum_{i=1}^{n} z_i \qquad z_i \in \{0, 1\},\ i = 1, \ldots, n$$

$$\lambda_0 \leq z_1 \qquad \lambda_i \leq z_i + z_{i+1},\ i = 1, \ldots, n-1$$

$$\lambda_n \leq z_n$$

Remark: $z_i = 1$ if $x \in [x_{i-1}, x_i]$.

Convex Combination (CC)

Data: $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$.

$$x = \sum_{i=0}^{n} \lambda_i x_i \qquad y = \sum_{i=0}^{n} \lambda_i y_i$$

$$1 = \sum_{i=0}^{n} \lambda_i \qquad \lambda_i \geq 0,\ i = 0, \ldots, n$$

$$1 = \sum_{i=1}^{n} z_i \qquad z_i \in \{0, 1\},\ i = 1, \ldots, n$$

$$\lambda_0 \leq z_1 \qquad \lambda_i \leq z_i + z_{i+1},\ i = 1, \ldots, n-1$$

$$\lambda_n \leq z_n$$

Remark: $z_i = 1$ if $x \in [x_{i-1}, x_i]$.

Incremental (INC)

Givens: $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$.

$$x = x_0 + \sum_{i=1}^{n} \delta_i \qquad\qquad y = y_0 + \sum_{i=1}^{n} \frac{(y_i - y_{i-1})}{(x_i - x_{i-1})}\delta_i$$

$$\delta_1 \leq (x_1 - x_0)$$

$$\delta_i \leq (x_i - x_{i-1})z_{i-1},\ i = 2, \ldots, n \qquad \delta_n \geq 0$$

$$\delta_i \geq (x_i - x_{i-1})z_i,\ i = 1, \ldots, n-1 \qquad z_i \in \{0, 1\},\ i = 1, \ldots, n-1$$

Remarks:

- If $z_i = 1$, then $z_j = 1$ for $j = 1, \ldots, i-1$.
- If $z_i = 1$, then $\delta_i = (x_i - x_{i-1})$.

Disaggregated Convex Combination (DCC)

Givens: $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$.

$$x = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} x_{i-1} + \lambda_i^{\mathrm{R}} x_i)$$

$$y = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} y_{i-1} + \lambda_i^{\mathrm{R}} y_i)$$

$$\lambda_i^{\mathrm{L}}, \lambda_i^{\mathrm{R}} \geq 0,\ i = 1, \ldots, n$$

$$z_i = \lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}},\ i = 1, \ldots, n$$

$$1 = \sum_{i=1}^{n} z_i$$

$$z_i \in \{0, 1\},\ i = 1, \ldots, n$$

Logarithmic Disaggregated Convex Combination (DLog)

Consists of a logarithmic encoding of the binary variables y_i, which correspond to intervals.

i	$\delta_3\delta_2\delta_1$
1	000
2	001
3	010
4	011
5	100
6	101
7	110
8	111

Let:

- $B_j^0 = \{i : \text{code of } i \text{ has value 0 at position } j\}$.
- $B_j^1 = \{i : \text{code of } i \text{ has value 1 at position } j\}$.

Example:

- $B_1^0 = \{1, 3, 5, 7\}$.
- $B_2^1 = \{3, 4, 7, 8\}$.

Logarithmic Disaggregated Convex Combination (DLog)

$$x = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} x_{i-1} + \lambda_i^{\mathrm{R}} x_i)$$

$$y = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} y_{i-1} + \lambda_i^{\mathrm{R}} y_i)$$

$$\lambda_i^{\mathrm{L}}, \lambda_i^{\mathrm{R}} \geq 0,\ i = 1, \ldots, n$$

$$\sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}}) = 1,$$

$$\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}} \leq \delta_j,\ \ i \in B_j^1,\ j = 1, \ldots, \lceil \log_2 n \rceil,$$

$$\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}} \leq 1 - \delta_j,\ \ i \in B_j^0,\ j = 1, \ldots, \lceil \log_2 n \rceil,$$

$$\delta_j \in \{0, 1\},\ j = 1, \ldots, \lceil \log_2 n \rceil$$

Logarithmic Disaggregated Convex Combination (DLog)

$$x = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} x_{i-1} + \lambda_i^{\mathrm{R}} x_i)$$

$$y = \sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} y_{i-1} + \lambda_i^{\mathrm{R}} y_i)$$

$$\lambda_i^{\mathrm{L}}, \lambda_i^{\mathrm{R}} \geq 0,\ i = 1, \ldots, n$$

$$\sum_{i=1}^{n} (\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}}) = 1,$$

$$\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}} \leq \delta_j,\ \ i \in B_j^1,\ j = 1, \ldots, \lceil \log_2 n \rceil,$$

$$\lambda_i^{\mathrm{L}} + \lambda_i^{\mathrm{R}} \leq 1 - \delta_j,\ \ i \in B_j^0,\ j = 1, \ldots, \lceil \log_2 n \rceil,$$

$$\delta_j \in \{0, 1\},\ j = 1, \ldots, \lceil \log_2 n \rceil$$

DCC e DLog

Remarks:

- DLog has the same number of continuous variables and constraints of DCC.
- However DLog needs a logarithmic number of binary variables.

SOS2

A set of variables, let us say $\{\lambda_0, \ldots, \lambda_n\}$, is SOS2 (Special Ordered Set of Variables Type 2) if:

1. At most two variables are positive.
2. If two variables are positive, then they are consecutive in the ordered set, let us say λ_i and λ_{i+1}.

Piecewise-Linear Model Based on SOS2

Givens: $\{(x_0, y_0),\ (x_1, y_1), \ldots, (x_n, y_n)\}$.

$$x = \sum_{i=0}^{n} \lambda_i x_i \qquad y = \sum_{i=0}^{n} \lambda_i y_i$$

$$1 = \sum_{i=0}^{n} \lambda_i \qquad \lambda_i \geq 0,\ i = 0, \ldots, n$$

$$\{\lambda_i\}_{i=0}^{n} \text{ is SOS2}$$

How Does SOS2 Work?

Implemented directly by the optimization solver:

- Suppose that $\{\lambda_0, \ldots, \lambda_n\}$ is a SOS2 set.
- Let $\{\tilde{\lambda}_0, \ldots, \tilde{\lambda}_n\}$ be the incumbent solution, in which $\tilde{\lambda}_{k_1}, \tilde{\lambda}_{k_2} > 0$ for $k_1, k_2 \in \{0, \ldots, n\}$, $k_1 < k_2$, and $k_2 - k_1 \geq 2$.

The infeasibility can be ruled out by "branching":

- Constraint $\lambda_0 = \cdots = \lambda_{k_1} = 0$ on the left branch.
- Constraint $\lambda_{k_1+2} = \cdots = \lambda_n = 0$ on the right branch.

How Does SOS2 Work?

Implemented directly by the optimization solver:

- Suppose that $\{\lambda_0, \dots, \lambda_n\}$ is a SOS2 set.
- Let $\{\tilde{\lambda}_0, \dots, \tilde{\lambda}_n\}$ be the incumbent solution, in which $\tilde{\lambda}_{k_1}, \tilde{\lambda}_{k_2} > 0$ for $k_1, k_2 \in \{0, \dots, n\}$, $k_1 < k_2$, and $k_2 - k_1 \geq 2$.

The infeasibility can be ruled out by "branching":

- Constraint $\lambda_0 = \cdots = \lambda_{k_1} = 0$ on the left branch.
- Constraint $\lambda_{k_1+2} = \cdots = \lambda_n = 0$ on the right branch.

Logarithmic Convex Combination (Log)

Remarks:

- Version of the CC model with a logarithmic number of variables and constraints.
- Needs an encoding corresponding to a Gray-Code.
- Complex structure of constraints, particularly in multidimensional domains.
- Requires a domain partitioning given by a J-1 triangulation.

Logarithmic Convex Combination (Log)

Remarks:

- Version of the CC model with a logarithmic number of variables and constraints.
- Needs an encoding corresponding to a Gray-Code.
- Complex structure of constraints, particularly in multidimensional domains.
- Requires a domain partitioning given by a J-1 triangulation.

Log Model: Example

Let $f : \mathbb{R} \to \mathbb{R}$ be a function with a piecewise-linear model:

- Set of breakpoints: $\mathcal{X} = \{x_0, x_1, \ldots, x_n\}$.
- Set of function values: $\mathcal{Y} = \{y_j = f(x_j) : x_j \in \mathcal{X}\}$.

Implementation:

- Set of domain intervals: $\mathcal{I} = \{i_1, \ldots, i_n\}$, such that
  - $i_1 = [x_0, x_1]$, $i_2 = [x_1, x_2]$, ..., $i_n = [x_{n-1}, x_n]$.
- $B : \mathcal{I} \to \{0, 1\}^{\log_2 |\mathcal{I}|}$ is a bijection defining a "Gray Code:"
  - $B(i)$ differs from $B(i+1)$ by just one bit.

Log Model: Example

Let $f : \mathbb{R} \to \mathbb{R}$ be a function with a piecewise-linear model:

- Set of breakpoints: $\mathcal{X} = \{x_0, x_1, \ldots, x_n\}$.
- Set of function values: $\mathcal{Y} = \{y_j = f(x_j) : x_j \in \mathcal{X}\}$.

Implementation:

- Set of domain intervals: $\mathcal{I} = \{i_1, \ldots, i_n\}$, such that
  - $i_1 = [x_0, x_1]$, $i_2 = [x_1, x_2]$, $\ldots$, $i_n = [x_{n-1}, x_n]$.
- $B : \mathcal{I} \to \{0, 1\}^{\log_2 |\mathcal{I}|}$ is a bijection defining a "Gray Code:"
  - $B(i)$ differs from $B(i+1)$ by just one bit.

Log Model: Example

Let $f : \mathbb{R} \to \mathbb{R}$ be a function with a piecewise-linear model:

- Set of breakpoints: $\mathcal{X} = \{x_0, x_1, \ldots, x_n\}$.
- Set of function values: $\mathcal{Y} = \{y_j = f(x_j) : x_j \in \mathcal{X}\}$.

Implementation:

- Set of domain intervals: $\mathcal{I} = \{i_1, \ldots, i_n\}$, such that
  - $i_1 = [x_0, x_1],\ i_2 = [x_1, x_2],\ \ldots,\ i_n = [x_{n-1}, x_n]$.
- $B : \mathcal{I} \to \{0, 1\}^{\log_2 |\mathcal{I}|}$ is a bijection defining a "Gray Code:"
  - $B(i)$ differs from $B(i + 1)$ by just one bit.

Log Model: Example

$$
\begin{array}{lcl}
B(i_1) = B([x_0, x_1]) & = & (0,0,0) \\
B(i_2) = B([x_1, x_2]) & = & (0,0,1) \\
B(i_3) = B([x_2, x_3]) & = & (0,1,1) \\
B(i_4) = B([x_3, x_4]) & = & (0,1,0) \\
B(i_5) = B([x_4, x_5]) & = & (1,1,0) \\
B(i_6) = B([x_5, x_6]) & = & (1,0,0) \\
B(i_7) = B([x_6, x_7]) & = & (1,0,1) \\
B(i_8) = B([x_7, x_8]) & = & (1,1,1)
\end{array}
$$

Log Model: Example

Let $J^+(B, l) \subseteq \mathcal{X}$ be the subset of breakpoints such that:

- for each $x \in J^+(B, l)$, the interval $I(x) \in \mathcal{I}$ to which it belongs, has value 1 at position l of the binary code $B(I(x))$.

Let $J^0(B, l) \subseteq \mathcal{X}$ be the subset of breakpoints such that:

- for each $x \in J^+(B, l)$, the interval $I(x) \in \mathcal{I}$ to which it belongs, has value 0 at position l of binary code $B(I(x))$.

Log Model: Example

Let $J^+(B,l) \subseteq \mathcal{X}$ be the subset of breakpoints such that:

- for each $x \in J^+(B,l)$, the interval $I(x) \in \mathcal{I}$ to which it belongs, has value 1 at position l of the binary code $B(I(x))$.

Let $J^0(B,l) \subseteq \mathcal{X}$ be the subset of breakpoints such that:

- for each $x \in J^+(B,l)$, the interval $I(x) \in \mathcal{I}$ to which it belongs, has value 0 at position l of binary code $B(I(x))$.

Log Model: Example

For $l = 1$:

$$J^+(B,1) = \{x_2, x_7, x_8\}$$
$$J^0(B,1) = \{x_0, x_4, x_5\}$$

For $l = 3$:

$$J^+(B,3) = \{x_5, x_6, x_7, x_8\}$$
$$J^0(B,3) = \{x_0, x_1, x_2, x_3\}$$

For $l = 2$:

$$J^+(B,2) = \{x_3, x_4, x_8\}$$
$$J^0(B,2) = \{x_0, x_1, x_6\}$$

Remark: This structure leads to a "branching scheme" compatible with SOS2.

Log Model: Example

For $l = 1$:

$$J^+(B,1) = \{x_2,\ x_7,\ x_8\}$$
$$J^0(B,1) = \{x_0,\ x_4,\ x_5\}$$

For $l = 2$:

$$J^+(B,2) = \{x_3,\ x_4,\ x_8\}$$
$$J^0(B,2) = \{x_0,\ x_1,\ x_6\}$$

For $l = 3$:

$$J^+(B,3) = \{x_5,\ x_6,\ x_7,\ x_8\}$$
$$J^0(B,3) = \{x_0,\ x_1,\ x_2,\ x_3\}$$

Remark: This structure leads to a "branching scheme" compatible with SOS2.

Log Model: Example

For $l = 1$:

$$J^+(B,1) = \{x_2,\ x_7,\ x_8\}$$
$$J^0(B,1) = \{x_0,\ x_4,\ x_5\}$$

For $l = 2$:

$$J^+(B,2) = \{x_3,\ x_4,\ x_8\}$$
$$J^0(B,2) = \{x_0,\ x_1,\ x_6\}$$

For $l = 3$:

$$J^+(B,3) = \{x_5,\ x_6,\ x_7,\ x_8\}$$
$$J^0(B,3) = \{x_0,\ x_1,\ x_2,\ x_3\}$$

Remark: This structure leads to a "branching scheme" compatible with SOS2.

Log Model: Example

For $l = 1$:

$$J^{+}(B,1) = \{x_2,\ x_7,\ x_8\}$$
$$J^{0}(B,1) = \{x_0,\ x_4,\ x_5\}$$

For $l = 2$:

$$J^{+}(B,2) = \{x_3,\ x_4,\ x_8\}$$
$$J^{0}(B,2) = \{x_0,\ x_1,\ x_6\}$$

For $l = 3$:

$$J^{+}(B,3) = \{x_5,\ x_6,\ x_7,\ x_8\}$$
$$J^{0}(B,3) = \{x_0,\ x_1,\ x_2,\ x_3\}$$

Remark: This structure leads to a "branching scheme" compatible with SOS2.

CC Model

$$f = \sum_{v \in \mathcal{V}} f(v)\lambda_v, \qquad x = \sum_{v \in \mathcal{V}} v\lambda_v, \tag{1}$$

$$1 = \sum_{v \in \mathcal{V}} \lambda_v, \tag{2}$$

$$\lambda_v \leq \sum_{P \in \mathcal{P}(v)} y_P, \qquad 1 = \sum_{P \in \mathcal{P}} y_P, \tag{3}$$

$$\lambda_v \geq 0, \; v \in \mathcal{V}, \tag{4}$$

$$y_P \in \{0, 1\}, \; P \in \mathcal{P} \tag{5}$$

Remark: The Log models offers a logarithmic representation of the equations (3) e (5).

Log Model

$$f = \sum_{v \in \mathcal{V}} f(v)\lambda_v, \tag{6}$$

$$x = \sum_{v \in \mathcal{V}} v\lambda_v, \tag{7}$$

$$1 = \sum_{v \in \mathcal{V}} \lambda_v, \tag{8}$$

$$\lambda_v \geq 0,\ v \in \mathcal{V}, \tag{9}$$

$$\sum_{v \in J^+(B,l)} \lambda_v \leq y_l,\ l \in \{1, \ldots, \lceil \log_2 |\mathcal{I}| \rceil\}, \tag{10}$$

$$\sum_{v \in J^0(B,l)} \lambda_v \leq (1 - y_l),\ l \in \{1, \ldots, \lceil \log_2 |\mathcal{I}| \rceil\}, \tag{11}$$

$$y_l \in \{0, 1\},\ l \in \{1, \ldots, \lceil \log_2 |\mathcal{I}| \rceil\}. \tag{12}$$

Piecewise-Linear Approximation: One Dimensional

- End!
- Thank you for your attention.