Spline Modeling of Cluster Tool Throughput

January 1999

David Ruppert
School of OR & IE
Cornell University

davidr@orie.cornell.edu
http://www.orie.cornell.edu/~davidr/ruppert.html

Spline Models of Cluster Tools

Aim:

Develop a simple mathematical model of a cluster tool:

- **Output**
  - Throughput
- **Input variables**
  - Pump time
  - Vent time
  - Robot speed
  - 1st step processing time
  - Alternative 1st step processing tiem
  - 2nd step processing tiem
  - etc.

Spline Models of Cluster Tools

Polynomial regression (univariate)

$$Y_i = \beta_0 + \beta_1 X_i + \cdots + \beta_p X_i^p + \text{error}$$

Spline regression (univariate)

$$Y_i = \beta_0 + \beta_1 X_i + \cdots + \beta_p X_i^p + \sum_{j=1}^{K} \beta_{p+j}(X_i - t_j)_+^p + \text{error}$$

- $t_1 < \ldots < t_K$ are fixed “knots”
- $u_+^p = u^p$ if $u > 0$ and equals 0 otherwise.
- $(x - t)_+^p$ is called a “truncated power function.”

Notation:

$$m(x; \boldsymbol{\beta}) = \beta_0 + \beta_1 X_i + \cdots + \beta_p X_i^p + \sum_{j=1}^{K} \beta_{p+j}(X_i - t_j)_+^p$$

Penalized Least Squares

$$\mathrm{PSS}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \Big\{ Y_i - m(x_i; \boldsymbol{\beta}) \Big\}^2 + \alpha \sum_{j=1}^{K} \beta_{p+j}^2$$

- $\widehat{\beta}(\alpha)$ minimizes PSS(β).
- Here $\alpha > 0$ is a penalty weight.
  - As $\alpha \to \infty$, the penalized spline converges to a pth degree polynomial fit.
  - As $\alpha \to 0$, the penalized spline converges to the OLS fitted spline.
- β_{p+j} is the jump in the pth derivative at the knot t_j. Thus, the penalty is on the $p+1$th derivative, interpreted as a generalized function.
- Solution is a ridge regression estimate.

Cross Validation (CV)

Goal: Chose α.

$m(x; \widehat{\boldsymbol{\beta}}(\alpha))$ is the fitted spline with penalty weight α.

$m_{-i}(x; \widehat{\boldsymbol{\beta}}(\alpha))$ is the fitted spline with penalty weight α and without using (X_i, Y_i).

$$\mathrm{CV}(\alpha) = \sum_{i=1}^{n} \left\{ Y_i - m_{-i}(X_i; \widehat{\boldsymbol{\beta}}(\alpha)) \right\}^2$$

Principle: Choose α that minimizes $\mathrm{CV}(\alpha)$.

Problem: CV is computationally intensive.

Solution: Use generalized cross validation (GCV).

Additive Models

- Suppose that there are J predictors, $X^{(1)}, \ldots, X^{(J)}$.
- An observation is $(X_i^{(1)}, \ldots, X_i^{(J)}, Y_i)$.
- An additive model is of the form

$$Y = \mu + m^{(1)}(X^{(1)}) + \cdots m^{(J)}(X^{(J)}) \ + \ \text{error}.$$

- The individual components (the $m^{(j)}$) are easily interpretable. This is the beauty of an additive model.
- The model is somewhat restrictive (all mixed partial derivatives are zero).

Interactions

- Suppose that we think that two predictor, say $X^{(1)}$ and $X^{(2)}$, may interact. Then we expand the additive model to

$$Y = \mu + m^{(1,2)}(X^{(1)}, X^{(2)}) + m^{(3)}(X^{(3)}) + \cdots m^{(J)}(X^{(J)}) + \text{error}. \tag{1}$$

- By adding enough interactions, one gets a full J-dimensional regression spline.
- Generally one wants something less than a full j-dimensional model because of the “curse of dimensionality.”

Spline Models of Cluster Tools

Example:

- All inputs except the first-step processing time are fixed.
- This is a univariate regression.

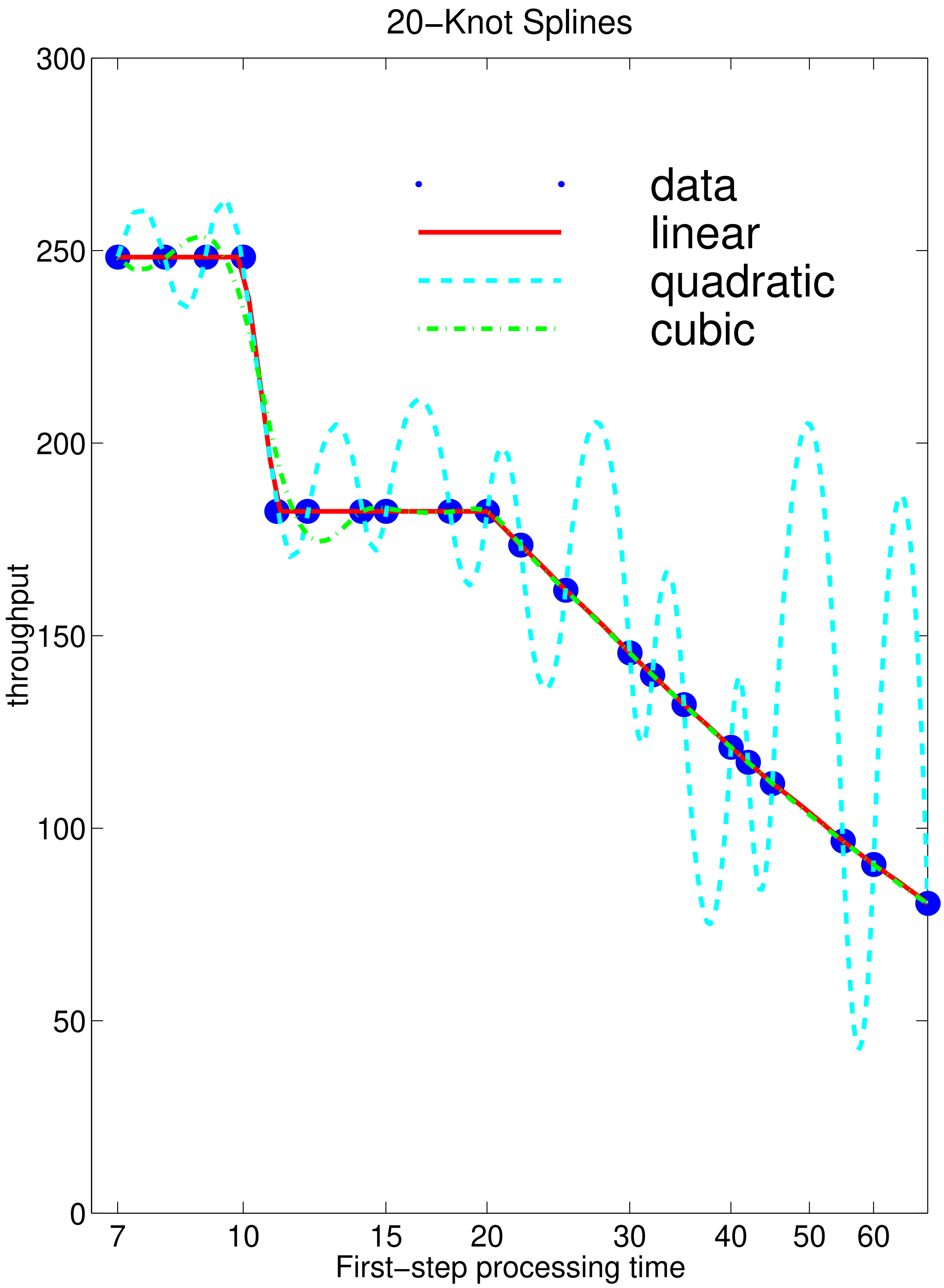

Figure 1: *Spline fits*