Functional Data Analysis
Chapters 16,17,18

Zhirong Yang

13 March 2007

Introduction

- 16. Functional linear models for functional responses
  - Generalize Ch.15 to a fully functional linear model where the response and covariate are all functions.
- 17. Derivatives and functional linear models
  - Two examples of the idea of a differential equation.
- 18. Differential equations and operators
  - A systematical look at how derivatives might be employed in modeling functional data.

Predicting Log Precipitation from Temperature

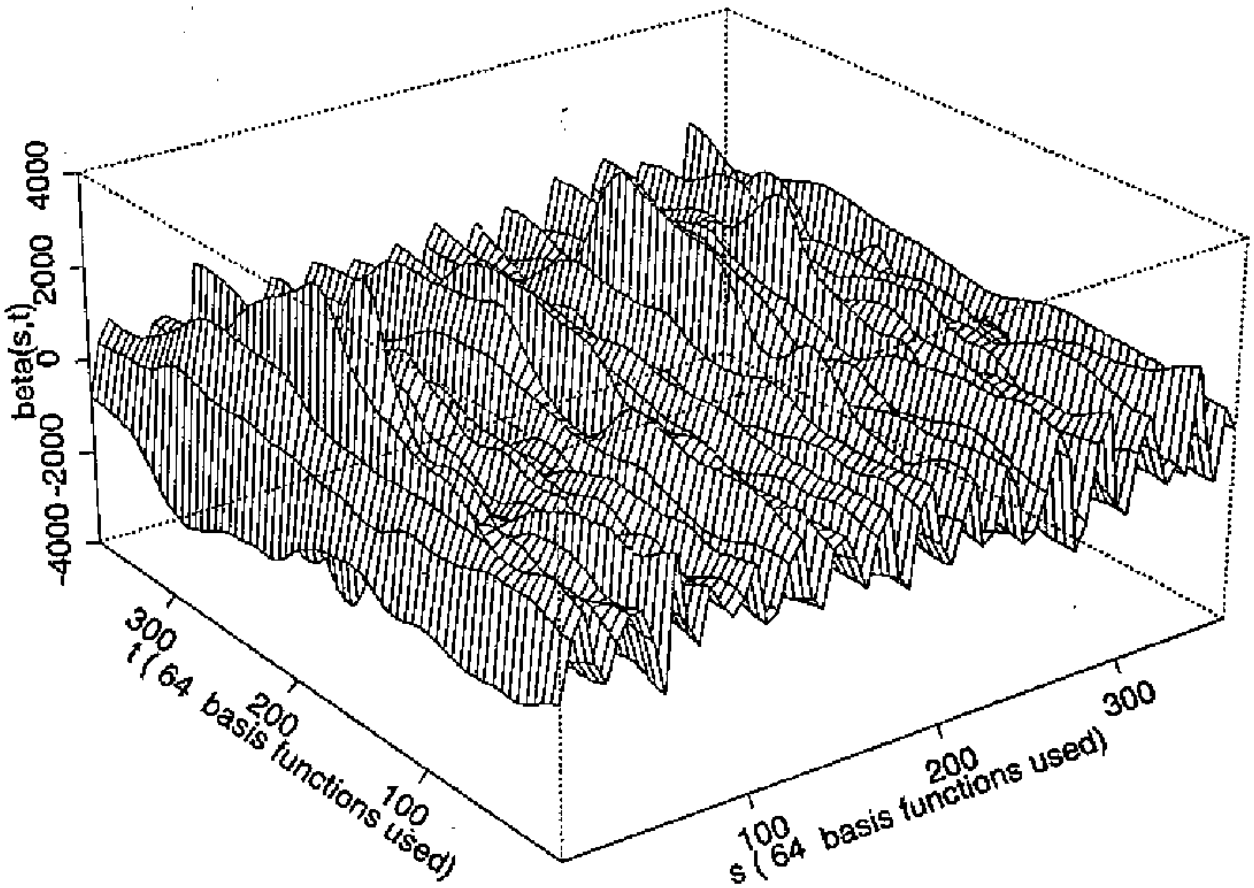
$$\text{LogPrec}_i(t) = \alpha(t) + \int_0^{365} \text{Temp}_i(s)\beta(s,t)ds + \epsilon_i(t)$$

$$\text{LMSSE}(\alpha, \beta) = \int \sum_{i=1}^{N} \left[\text{LogPrec}_i(t) - \alpha(t) - \int \text{Temp}_i(s)\beta(s,t)ds \right]^2 dt$$

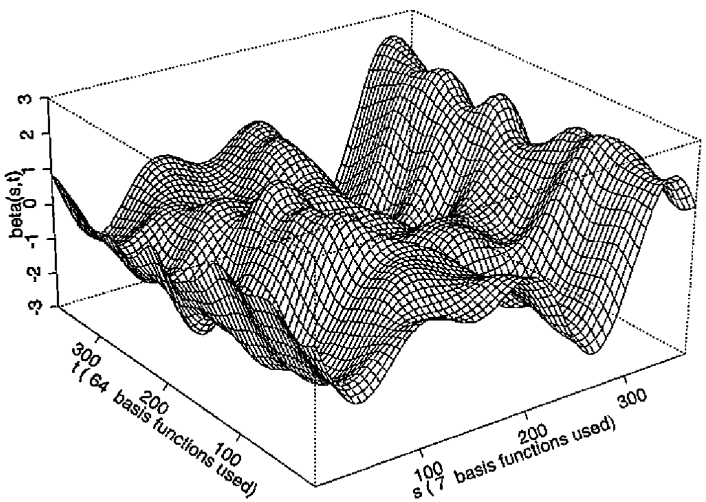
$$\beta(s,t) = \sum_{k=1}^{K_1} \sum_{l=1}^{K_2} b_{kl}\eta_k(s)\theta_l(t) = \boldsymbol{\eta}(s)'\mathbf{B}\boldsymbol{\theta}(t)$$

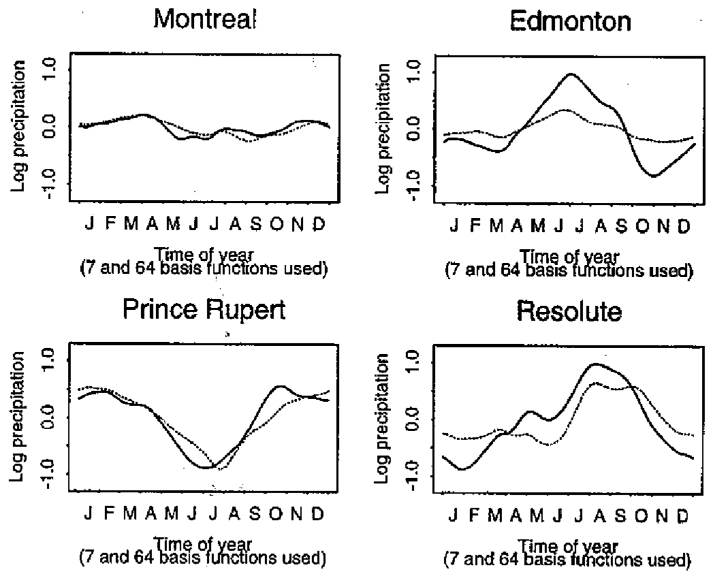
$$\alpha(t) = \sum_{l=1}^{K_2} a_l\theta_l(t) = \boldsymbol{\theta}'(t)\mathbf{a}$$

Fitting The Model without Regularization

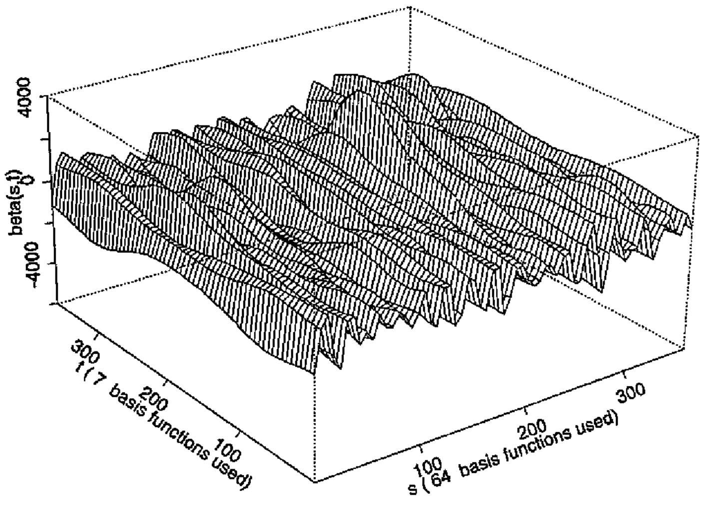


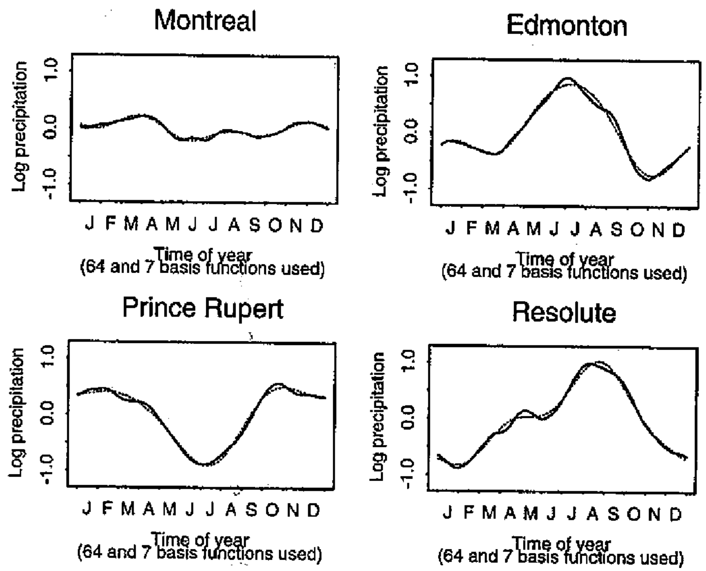
Restricting The Basis $\boldsymbol{\eta}(s)$



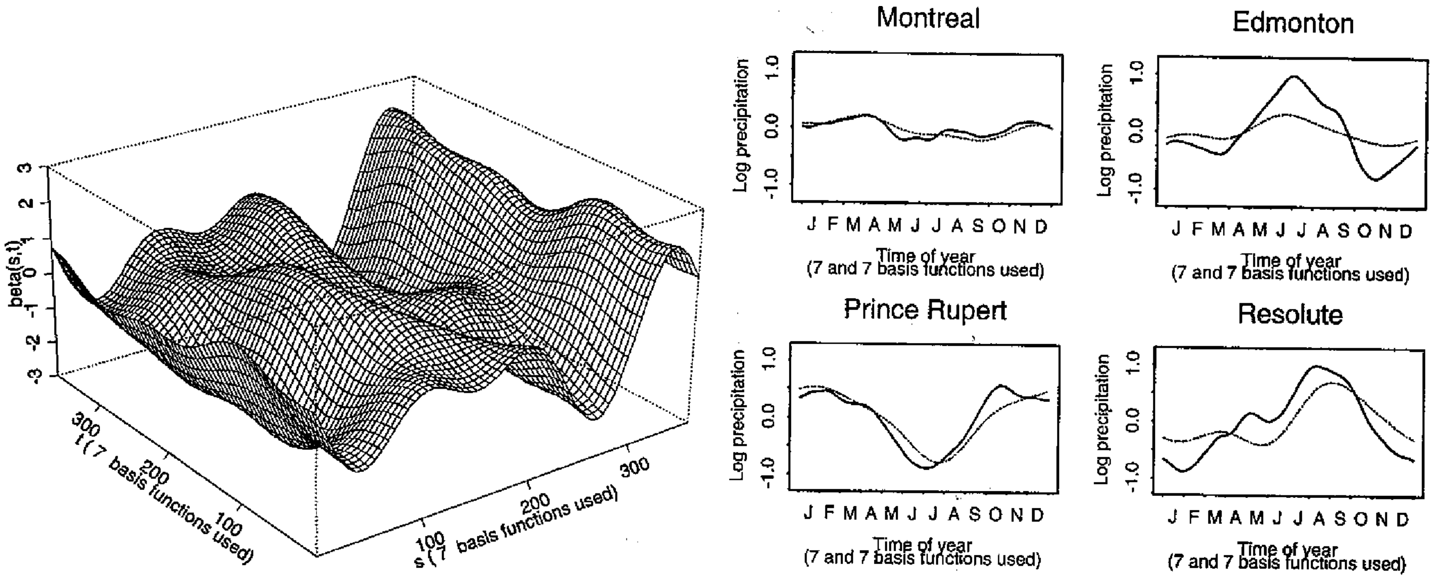


Restricting The Basis $\boldsymbol{\theta}(t)$

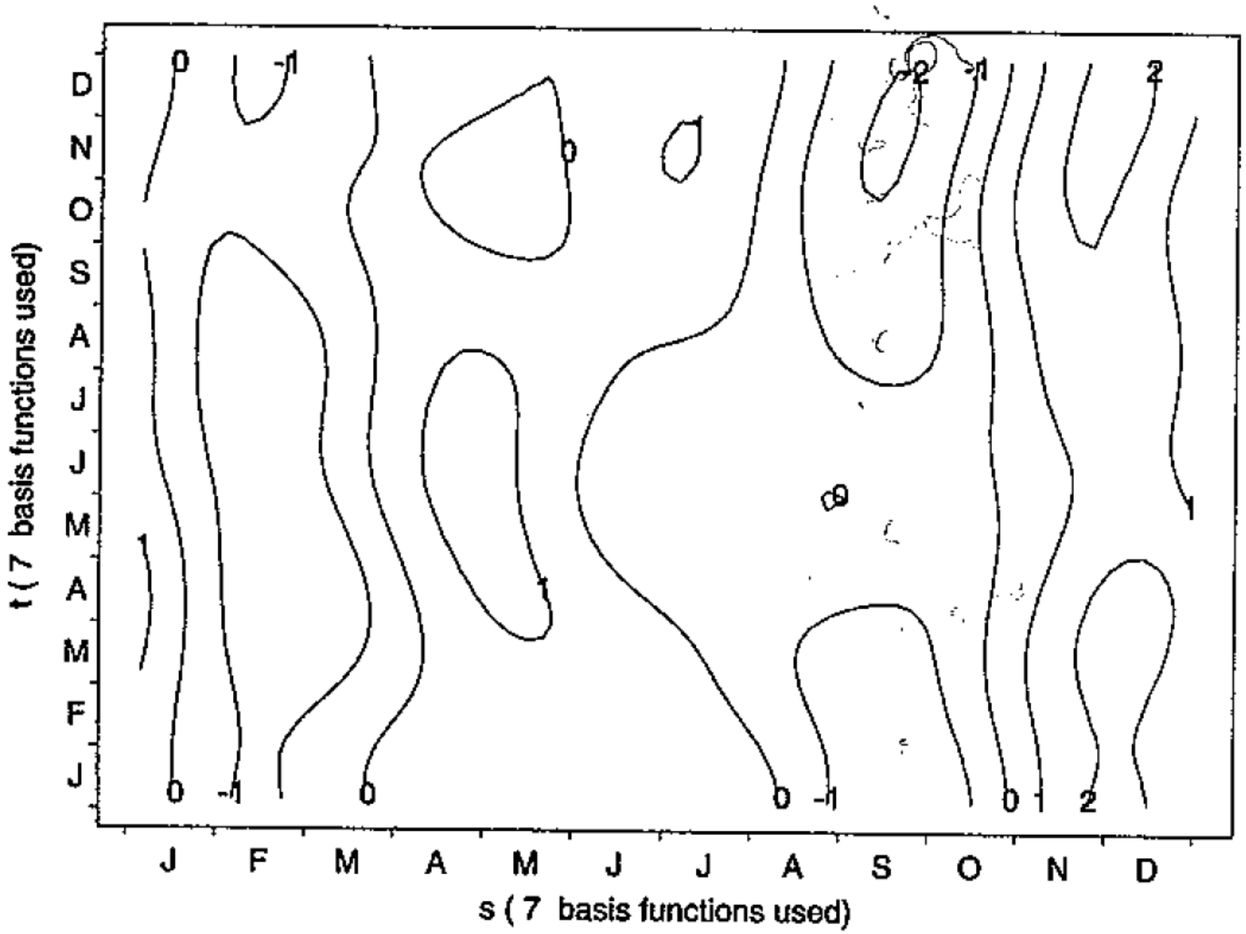




Restricting Both Bases



Restricting Both Bases (cont.)



Assessing Goodness of Fit

$$R^2(t) = 1 - \sum_i \{\hat{y}_i(t) - y_i(t)\}^2 \Big/ \sum_i \{y_i(t) - \bar{y}(t)\}^2$$

$$R_i^2 = 1 - \int \{\hat{y}_i(t) - y_i(t)\}^2 \, dt \Big/ \int \{y_i(t) - \bar{y}(t)\}^2 \, dt$$

$$\text{FRATIO}(t) = \frac{\sum_i \{\hat{y}_i(t) - \bar{y}(t)\}^2/(K_0 - 1)}{\sum_i \{y_i(t) - \hat{y}_i(t)\}^2/(n - K_0)}$$

$$\hat{y}_i(t) - \bar{y}(t) = \sum_{j=1}^{J_0} C_{ij} \left(\sum_{k=1}^{K_0} B_{ij} \theta_k(t) \right) = \sum_{j=1}^{J_0} C_{ij} \theta_j(t)$$

Computational Details

$$\begin{aligned}\mathbf{y}^*(t) &= \int \mathbf{z}^*(s)\beta(s,t)ds + \boldsymbol{\epsilon}(t) \\ &= \int \mathbf{z}^*(s)\boldsymbol{\theta}'(s)\mathbf{B}\boldsymbol{\eta}(t)ds + \boldsymbol{\epsilon}(t) \\ &= \mathbf{Z}^*\mathbf{B}\boldsymbol{\eta}(t) + \boldsymbol{\epsilon}(t)\end{aligned}$$

$$\mathbf{Z}^* = \int \mathbf{z}^*(s)\boldsymbol{\theta}'(s)ds$$

$$\mathbf{Z}^{*\prime}\mathbf{Z}^*\mathbf{B}\int \boldsymbol{\eta}(t)\boldsymbol{\eta}'(t)dt = \mathbf{Z}^{*\prime}\int y(t)\boldsymbol{\eta}'(t)dt$$

Computational Details (with Regularization)

$$
\begin{aligned}
\mathrm{PEN}_s(\beta) &= \int\int \left[L_s\beta(s,t)\right]^2 ds dt \\
&= \int\int \left[L_s\boldsymbol{\theta}'(s)\mathbf{B}\boldsymbol{\eta}(t)\right]\left[L_s\boldsymbol{\theta}'(s)\mathbf{B}\boldsymbol{\eta}(t)\right]' ds dt \\
&= \int\int \left[L_s\boldsymbol{\theta}'(s)\right]\mathbf{B}\boldsymbol{\eta}(t)\boldsymbol{\eta}'(t)\mathbf{B}'\left[L_s\boldsymbol{\theta}(s)\right] ds dt \\
&= \int \mathrm{trace}\left[\mathbf{B}\boldsymbol{\eta}(t)\boldsymbol{\eta}'(t)\mathbf{B}'\mathbf{R}\right] dt \\
&= \mathrm{trace}\left[\mathbf{B}'\mathbf{R}\mathbf{B}\mathbf{J}_{\eta\eta}\right]
\end{aligned}
$$

$$
\mathbf{R} = \int \left[L_s\boldsymbol{\theta}(s)\right]\left[L_s\boldsymbol{\theta}'(s)\right] ds \qquad \mathbf{J}_{\eta\eta} = \int \boldsymbol{\eta}(t)\boldsymbol{\eta}'(t) dt
$$

Computational Details (with Regularization)

$$\text{PEN}_t(\beta) = \text{trace}\left[\mathbf{B}'\mathbf{J}_{\theta\theta}\mathbf{S}\mathbf{B}\right]$$

$$\mathbf{S} = \int \left[L_t \boldsymbol{\eta}(t)\right] \left[L_t \boldsymbol{\theta}'(t)\right] dt \qquad \mathbf{J}_{\theta\theta} = \int \boldsymbol{\theta}(s)\boldsymbol{\theta}'(s) ds$$

$$\mathbf{Z}^{*\prime}\mathbf{Z}^{*}\mathbf{B}\mathbf{J}_{\eta\eta} + \lambda_s \mathbf{R}\mathbf{B}\mathbf{J}_{\eta\eta} + \lambda_t \mathbf{J}_{\theta\theta}\mathbf{B}\mathbf{S} = \mathbf{Z}^{*\prime} \int \mathbf{y}\boldsymbol{\eta}'$$

The General Case

$$\beta(s,t) = \sum_{k}^{K_\beta} b_k \theta_k(s,t) = \boldsymbol{\theta}'(s,t)\mathbf{b}$$

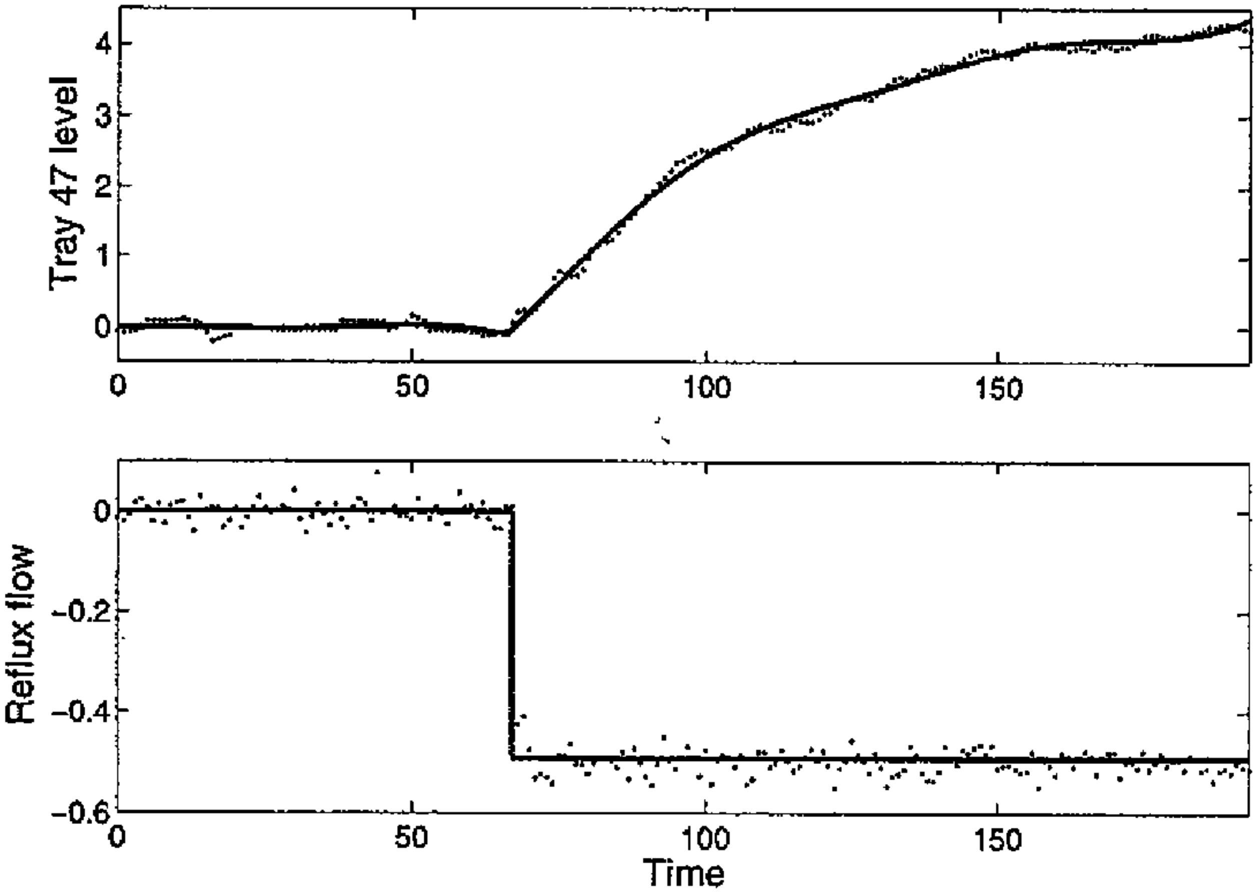
$$y_i(t) = \int_{\Omega_t} z_i(s,t)\beta(s,t)ds + \epsilon_i(t) = \int_{\Omega_t} z_i(s,t)\boldsymbol{\theta}'(s,t)\mathbf{b}ds + \epsilon_i(t)$$

$$z_{ik}^*(t) = \int_{\Omega_t} z_i(s,t)\theta_k'(s,t)ds$$

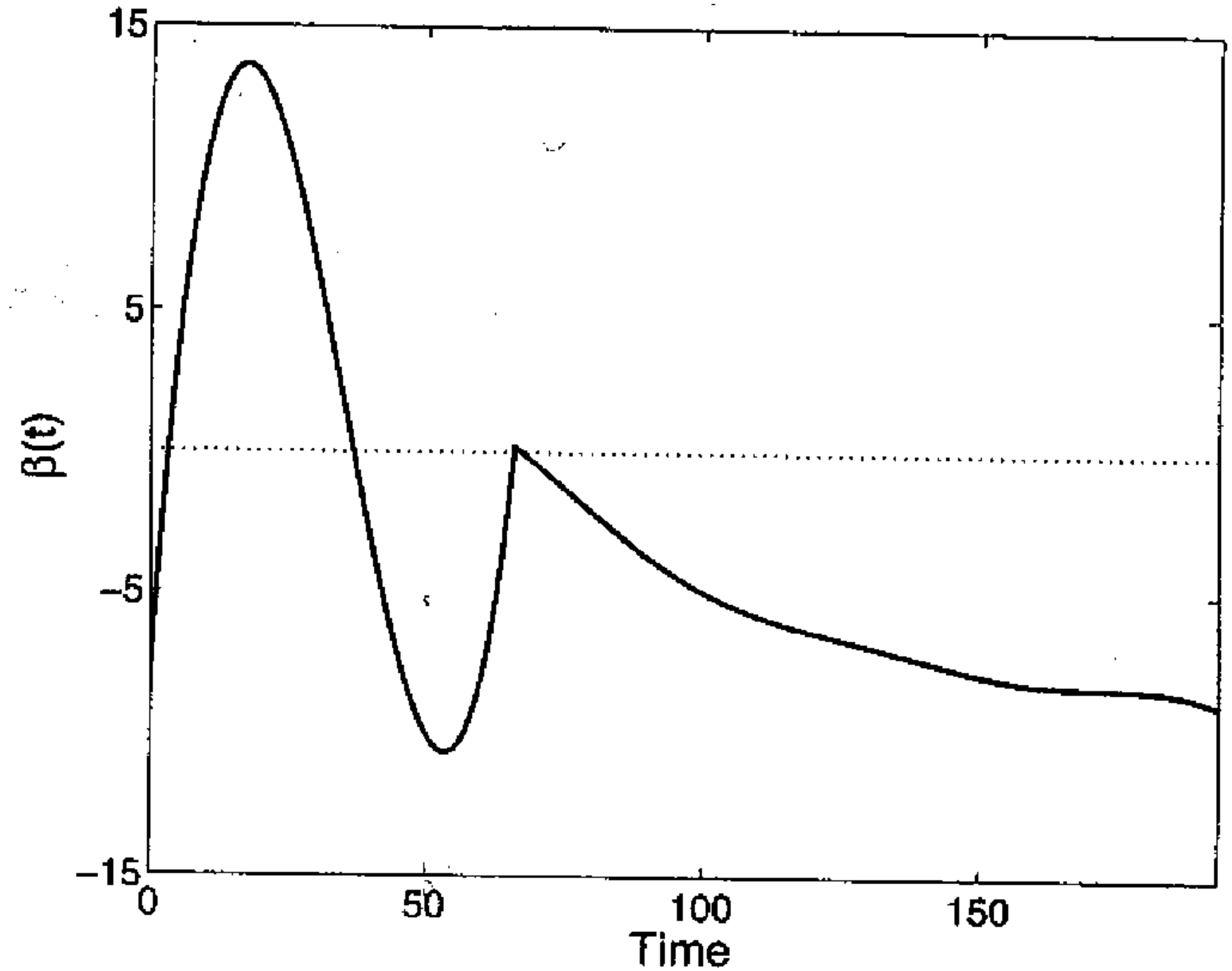
$$\left[\int \mathbf{Z}^{*\prime}\mathbf{Z}^* + \lambda_s\mathbf{R} + \lambda_t\mathbf{S}\right]\hat{\mathbf{b}} = \int \mathbf{Z}^{*\prime}\mathbf{y}$$

$$\mathbf{R} = \int\int_{\Omega_t} \left[L_s\boldsymbol{\theta}(s,t)\right]\left[L_s\boldsymbol{\theta}(s,t)\right]' dsdt \quad \mathbf{S} = \int\int_{\Omega_t} \left[L_t\boldsymbol{\theta}(s,t)\right]\left[L_t\boldsymbol{\theta}(s,t)\right]' dsdt$$

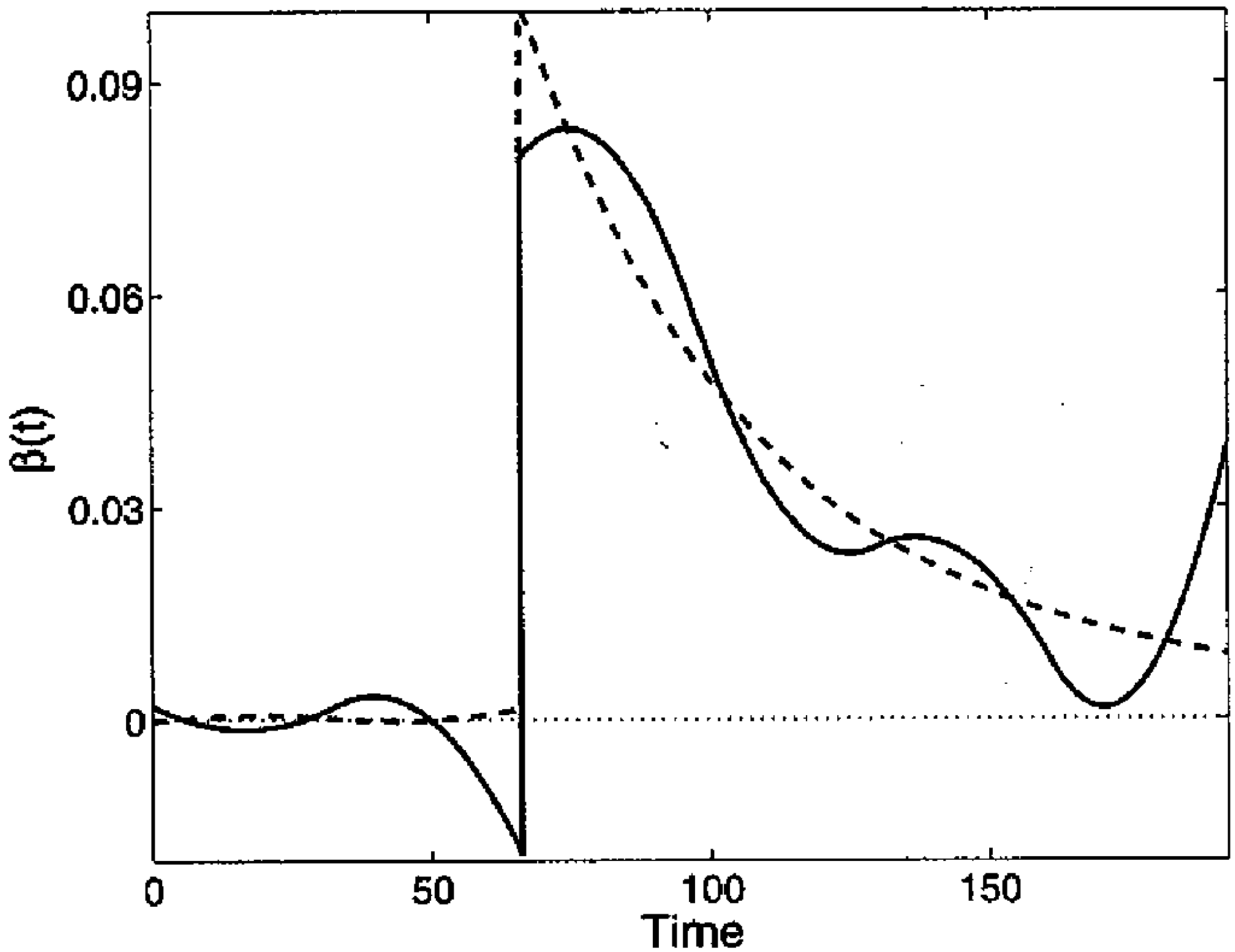
The Oil Refinery Data



$$\mathbf{Tray}(t) = \mathbf{Reflux}(t)\beta(t) + \epsilon(t)$$



$$D\mathbf{Tray}(t) = -\beta_1\mathbf{Tray}(t) + \beta_2(t)\mathbf{Reflux}(t) + \epsilon(t)$$



$$D\textbf{Tray}(t) = -\beta_1 \textbf{Tray}(t) + \beta_2(t)\textbf{Reflux}(t) + \epsilon(t)$$

Let $y(t) = \text{Tray}(t)$, $x(t) = \text{Reflux}(t)$. We get

$$y(t) = e^{-\beta_1 t}\left[y(0) - \frac{\beta_2}{\beta_1}\int_0^t e^{\beta_1 s} u(s) ds\right],$$

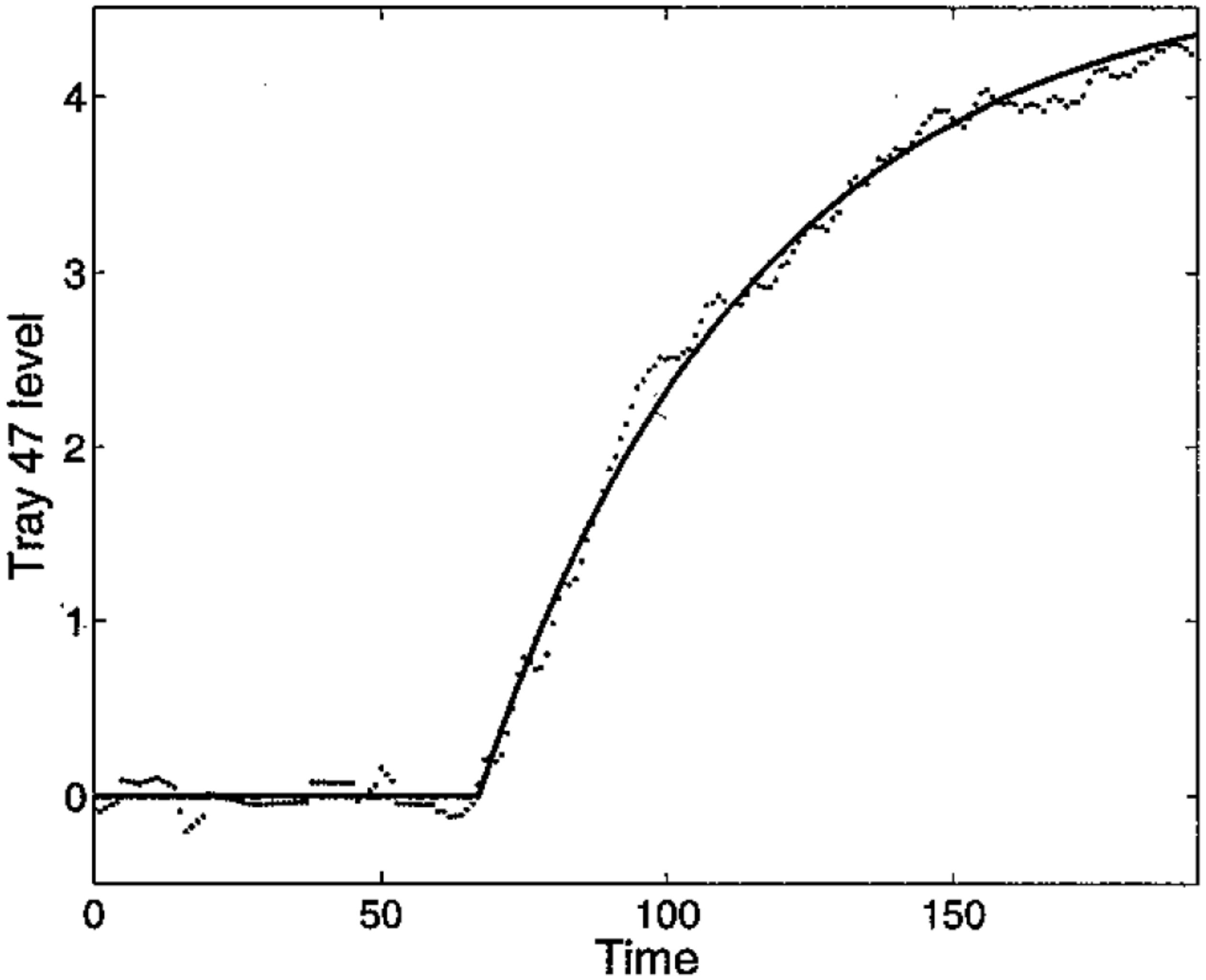
which can be simplified to

$$y(t) = 0.4924\frac{\beta_2}{\beta_1}\left[1 - e^{-\beta_1(t-67)}\right],\ t \geq 67,\ \text{and 0 otherwise}$$

by setting

$$y(0) = 0, u(t) = 0, t \leq 67,\ \text{and}\ u(t) = -0.4924,\ t > 67.$$

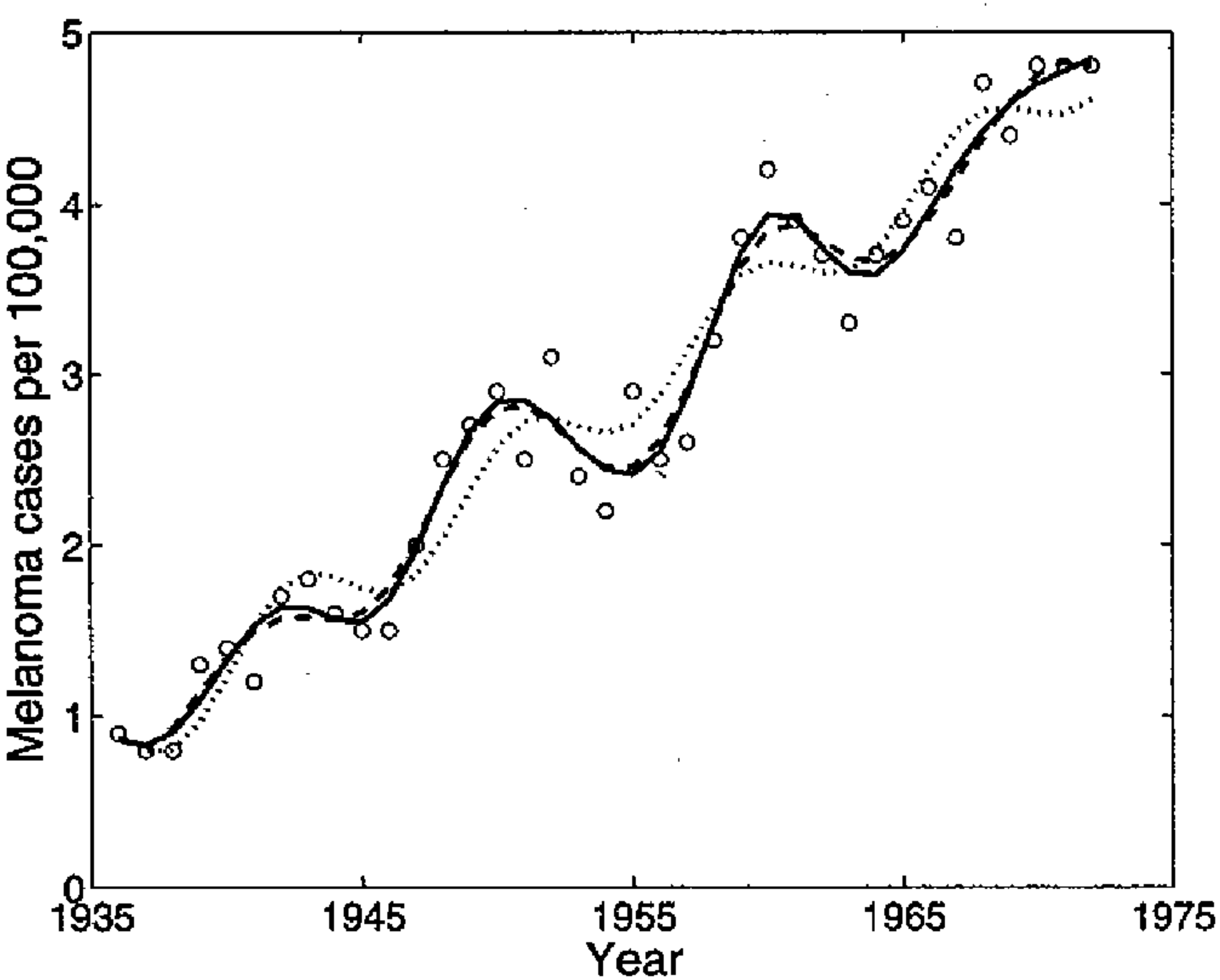
$$D\mathbf{Tray}(t) = -\beta_1 \mathbf{Tray}(t) + \beta_2(t)\mathbf{Reflux}(t) + \epsilon(t)$$



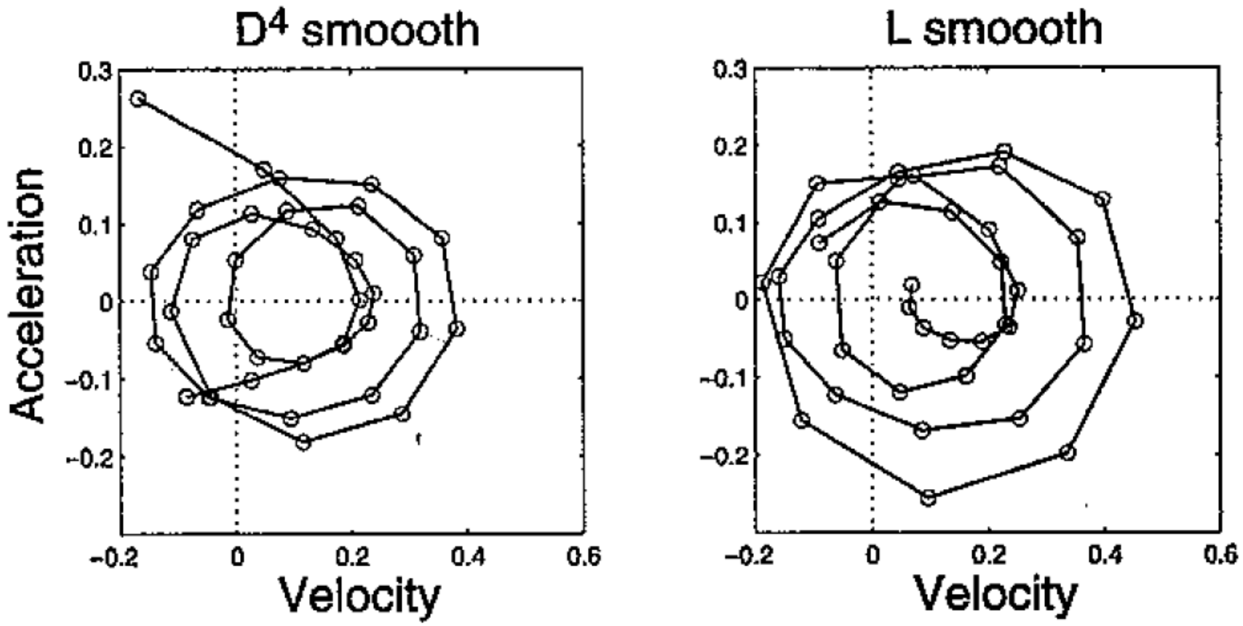
The Melanoma Data

- $D^4 x = -\beta_1 D^2 x - \beta_2 D^3 x$
- Algorithm steps:
  1. Do smoothing by using penalty of order D^4
  2. Compute the derivatives of the smooth up to order four
  3. Estimate β_1 and β_2 by e.g. regression
  4. Solve $Lx = \beta_1 D^2 x + \beta_2 D^3 x + D^4 x = 0$
  5. Do smoothing again by using penalty defined by the linear differential operator.
  6. Go back to step 2 if not converged.

The Melanoma Data (cont.)



The Melanoma Data (cont.)



Exploring A Simple Linear Differential Equation

- $Dx(t) = -\beta x(t) + \epsilon(t)$ (*Homogeneous* Equation)

$$x(t) = Ce^{-\beta t}$$

- $Dx(t) = -\beta x(t) + \alpha u(t) + \epsilon(t)$

$$x(t) = Ce^{-\beta t} + \alpha \int_0^t e^{-\beta(t-s)} u(s) ds$$

Let $u(t) = 0$ for $0 \leq t \leq 1$, $u(t) = 1$ for $t > 1$; $x(0) = C = 1$.

$$x(t) = \begin{cases} e^{-\beta t} & 0 \leq t \leq 1 \\ Ce^{-\beta t} + (\alpha/\beta)\left[1 - e^{-\beta(t-1)}\right] & t > 1 \end{cases}$$

Exploring A Simple Linear Differential Equation (cont.)

- Rearrange the nonhomogeneous equation

$$Lx(t) = \beta x(t) + Dx(t) - \alpha u(t) - \epsilon(t)$$

- Function x is a solution of the original equation when $\epsilon = 0$ if and only if $Lx = 0$.

- We call $L = \beta I + D$ a *linear differential operator*.

Beyond The Constant Coefficient First-Order Linear Equation

- Nonconstant coefficient

$$Dx(t) = -\beta(t)x(t) + \alpha(t)u(t) + \epsilon(t)$$

- Higher order equations

$$D^m x(t) = -\sum_{j=0}^{m-1} \beta_j(t) D^j x(t) + \alpha(t)u(t) + \epsilon(t)$$

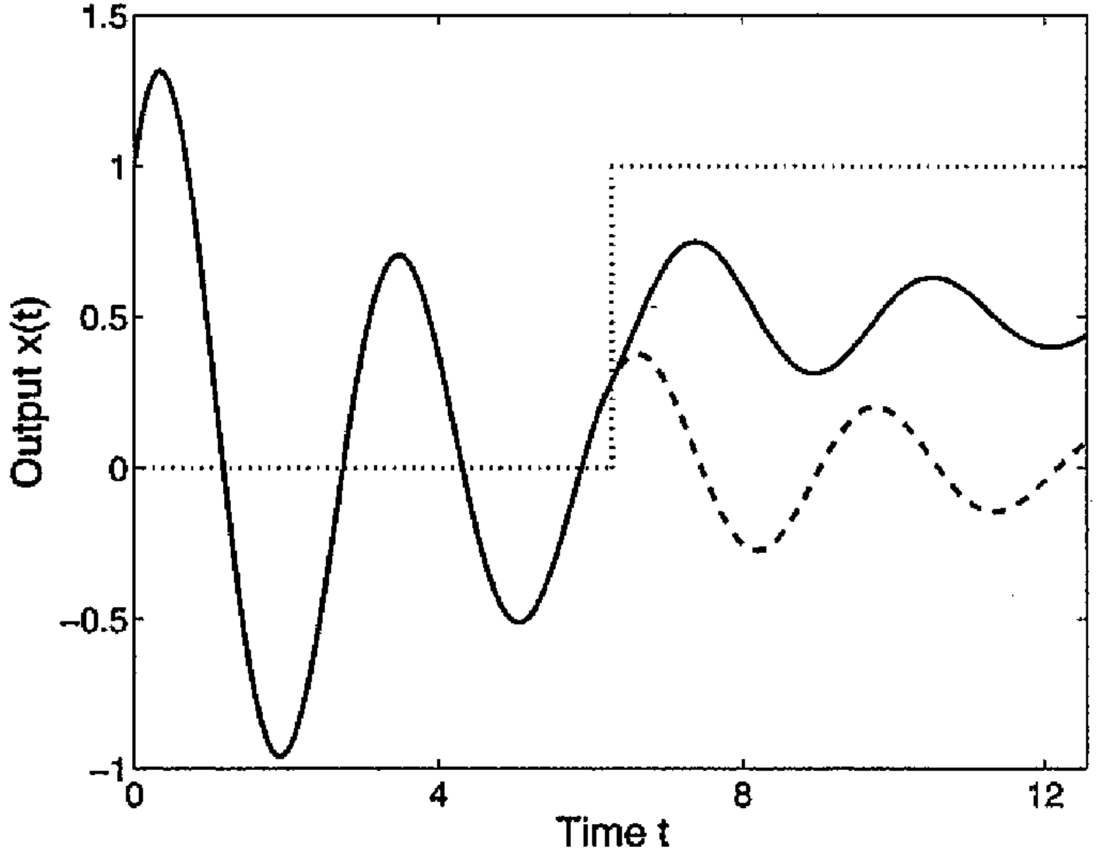
- Systems of equations

$$Dx(t) = -\beta_x(t)x(t) + \alpha_x(t)u(t)$$
$$Du(t) = -\beta_u(t)u(t) + \alpha_u(t)x(t)$$

- Beyond linearity $Dx(t) = f\left[t, x(t), u(t)\right]$

$$D^2x(t) = -4.04x(t) - 0.4Dx(t) + 2u(t)$$

$$x(t) = e^{-0.2t}\left[\sin(2t) + \cos(2t)\right]$$



Some Applications of Linear Differential Equations and Operators

- Newton's third law

$$s(t) = s_0 + v_0 t + M^{-1} \int_0^t \int_0^u F(z) dz du$$

- The gross domestic product data

$$Lx = \beta x + Dx$$

Differential Operators to Regularize Models

$$\text{PENSSE}_\lambda(\hat{x}) = \frac{1}{n}\sum_j [y_j - \hat{x}(t_j)]^2 + \lambda \int (L\hat{x})^2(t)dt$$

It follows that we can use a relatively large value of the smoothing parameter λ, leading to lower variance, without introducing excessive bias.

Differential Operators to Partition Variation

- The functions ξ_j that satisfy $L\xi_j = 0$ forms a basis of the null space of L, i.e. ker$L$.
- Linear differntial operators make use of the information in derivatives while projection operators do not.

$$x(t) = c_0 + \sum_{k=1}^{\infty} [c_{2k-1}\sin(2\pi kt) + c_{2k}\cos(2\pi kt)]$$

$$Qx(t) = \sum_{k=2}^{\infty} [c_{2k-1}\sin(2\pi kt) + c_{2k}\cos(2\pi kt)]$$

$$\begin{aligned} Lx(t) &= 4\pi^2 Dx(t) + D^3x(t) \\ &= \sum_{k=2}^{\infty} 8\pi^3 k(k^2-1)\,[-c_{2k-1}\sin(2\pi kt) + c_{2k}\cos(2\pi kt)] \end{aligned}$$

Finding A Linear Differential Operator That Annihilates Known Functions

$$x(t) = A(t)\left[c_1 \sin(\omega t) + c_2 \cos(\omega t)\right]$$

$$Lx = \beta_0 x + \beta_1 Dx + D^2 x$$

$$\boldsymbol{\beta}(t) = \begin{bmatrix} \beta_0(t) \\ \beta_1(t) \end{bmatrix} \qquad \boldsymbol{\xi}(t) = \begin{bmatrix} A(t)\sin(\omega t) \\ A(t)\cos(\omega t) \end{bmatrix} = \begin{bmatrix} AS \\ AC \end{bmatrix}$$

$$\beta_0 \boldsymbol{\xi} + \beta_1 D\boldsymbol{\xi} = -D^2 \boldsymbol{\xi} \quad \text{or} \quad \begin{bmatrix} \boldsymbol{\xi} & D\boldsymbol{\xi} \end{bmatrix} \boldsymbol{\beta} = -D^2 \boldsymbol{\xi}$$

Wronskian matrix

$$\mathbf{W}(t) = \begin{bmatrix} \boldsymbol{\xi} & D\boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} AS & (DA)S + \omega AC \\ AC & (DA)C - \omega AS \end{bmatrix}$$

Finding A Linear Differential Operator That Annihilates Known Functions (cont.)

Wronskian

$$|\mathbf{W}| = AS[(DA)C - \omega AS] - AC[(DA)S + \omega AC] = -\omega A^2$$

$$\boldsymbol{\beta} = -\mathbf{W}^{-1}D^2\boldsymbol{\xi} = \begin{bmatrix} \omega^2 + 2(DA/A^2)^2 - D^2A/A \\ -2DA/A \end{bmatrix}$$

$$Lx = [\omega^2 + 2(DA/A)^2 - D^2A/A]x - 2[(DA)/A](Dx) + D^2x$$

Finding A Linear Differential Operator That Annihilates Known Functions (cont.)

- If A is constant,

$$L = \omega^2 I + D^2.$$

- If $A(t) = e^{-\lambda t}$,

$$\boldsymbol{\beta} = \begin{bmatrix} \omega^2 + \lambda \\ 2\lambda \end{bmatrix} \quad \text{or} \quad Lx = (\omega^2 + \lambda^2)x + 2\lambda Dx + D^2 x.$$

- For high-order equations,

$$\mathbf{W}(t) = \begin{bmatrix} \boldsymbol{\xi}(t) & D\boldsymbol{\xi}(t) & \ldots & D^{-1}\boldsymbol{\xi}(t) \end{bmatrix}$$

$$\mathbf{W}(t)\boldsymbol{\beta}(t) = -D^m \boldsymbol{\xi}(t),$$

which is solved by numerical differential equation solvers.

Additional Constraints Needed to Define A Solution

Any specific solution of $Lx = 0$ requires m additional pieces of information about x.

Initial Operator
$$B_0 x = \begin{bmatrix} x(0) \\ Dx(0) \\ \vdots \\ D^{m-1}x(0) \end{bmatrix}$$

Boundary Operator
$$B_B x = \begin{bmatrix} x(0) \\ x(T) \\ \vdots \\ D^{(m-2)/2}x(0) \\ D^{(m-2)/2}x(T) \end{bmatrix}$$

Additional Constraints (cont.)

$$\text{Perodic Operator } B_P x = \begin{bmatrix} x(T) - x(0) \\ Dx(T) - Dx(0) \\ \vdots \\ D^{m-1}x(T) - D^{m-1}x(0) \end{bmatrix}$$

$$\text{Integral Operator } B_I x = \begin{bmatrix} \int \xi_1(t)x(t)dx \\ \int \xi_2(t)x(t)dx \\ \vdots \\ \int \xi_m(t)x(t)dx \end{bmatrix}$$

The inner products defined by operators L and B

$$\langle x, y \rangle_{B,L} = (Bx)'(By) + \int (Lx)(t)(Ly)(t)dt$$

$$\|x\|^2_{B,L} = (Bx)'(Bx) + \int (Lx)^2(t)dt$$

$x = z + e$ where $z \in \ker L$ and $e \in \ker B$

$$\|x\|^2_{B,L} = \|z\|^2_B + \|e\|^2_L$$