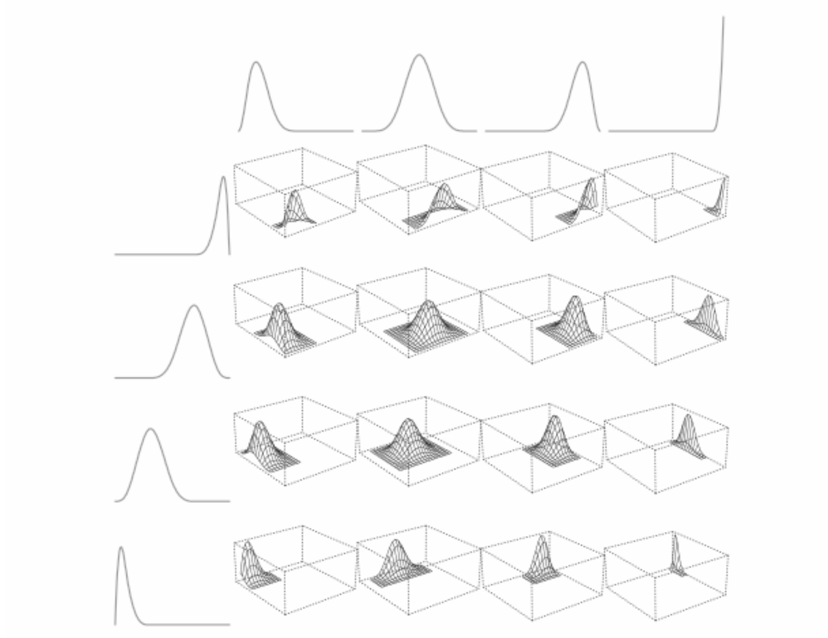
Multidimensional Splines

- Suppose $X \in R^2$
- We have a separate basis of functions for representing functions of X1 and X2

$$h_{1k}(X_1),\ k = 1,\ldots,M_1$$
$$h_{2k}(X_2),\ k = 1,\ldots,M_2$$

- The M1xM2 dimensional tensor product basis is

$$g_{jk} = h_{1j}(X_1)h_{2k}(X_2),\ \ j = 1,\ldots,M_1,\ \ k = 1,\ldots,M_2$$



Tensor Product

- Or direct product, or Krondecker product…

  -A is an m by n matrix, B is a p by q matrix

$$C = A \otimes B$$

  -C is an mp by nq matrix with elements

$$c_{\alpha\beta} = a_{ij} b_{kl}$$

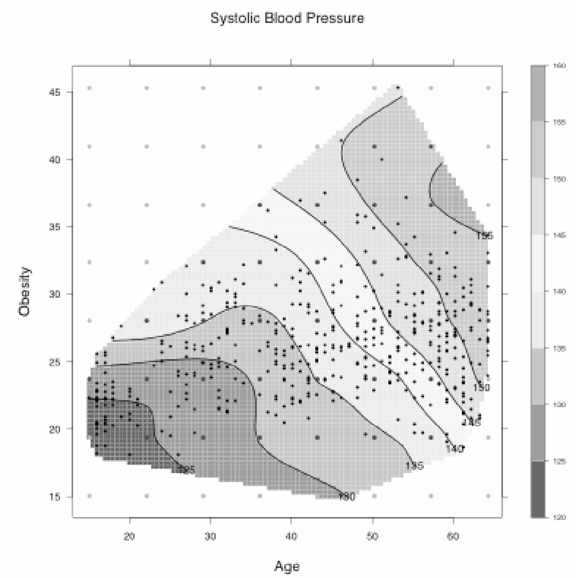
$$\alpha = p(i-1)+k$$

$$\beta = q(j-1)+l$$

2D Function Basis

$$g(X) = \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} \theta_{jk} g_{jk}(X)$$

- Can be generalized to higher dimensions
- Dimension of the basis grows exponentially in the number of coordinates (curse of dimensionality)
- MARS (Ch 9) is a greedy algorithm for including only the basis functions deemed necessary by least squares



Questions

- 1. In Figure 5.11 The tensor product spline is said to have some "spurious"boundaries. From the figure they seem a little more serious than that, especially the cut in the far left of the figure. This seems like a similair effect to the "flapping" described earlier in the chapter at the boundary conditions. Does this mean that we have to deal with the boundaries more carefully when using tensor products?
  - Ashish

Questions

- 2. Smoothing as defined as the penalty component in the optimization seems to be "independant" from the the choice of the type of basis. Could using more complex basis function imply more creaticvity is required in thesmoothing function, like using second/tensor derivatives?
- 3. p140 J(F) finds one lambda over both X1, X2 derivate constraints. Could we generalize further by finding a separate lambda for each xi instead? What would the effects on the final function be?
  - Ashish

Higher Dimensional Smoothing Splines

- Suppose we have pairs $y_i, x_i, \; x_i \in R^d$
- Want to find d-dimensional regression function f(x)
- Solve

$$\min_f \sum_{i=1}^{N} \{y_i - f(x_i)\}^2 + \lambda J[f]$$

- J is a penalty functional for stabilizing f in R^d

2D Roughness Penalty

- Generalization of 1D penalty

$$J[f] = \iint_{R^2} [(\frac{\partial^2 f(x)}{\partial x_1^2})^2 + 2(\frac{\partial^2 f(x)}{\partial x_1 \partial x_2})^2 + (\frac{\partial^2 f(x)}{\partial x_2^2})^2] dx_1 dx_2$$

- Yields thin plate spline (smooth 2D surface)

Thin Plate Spline

- Properties in common with 1D cubic smoothing spline
  - As $\lambda \to 0$, solution approaches interpolating fn
  - As $\lambda \to \infty$, solution approaches least squares plane
  - For intermediate values of λ, solution is a linear expansion of basis functions with coefficients obtained by a generalized form of ridge regression

Thin Plate Spline

- Solution has form

$$f(x) = \beta_0 + \beta^T x + \sum_{j=1}^{N} \alpha_j h_j(x)$$

$$h_j(x) = \eta(\|x - x_j\|)$$

$$\eta(z) = z^2 \log z^2$$

- hj are radial basis functions
- Can be generalized to arbitrary dimension

Computational Speed-ups

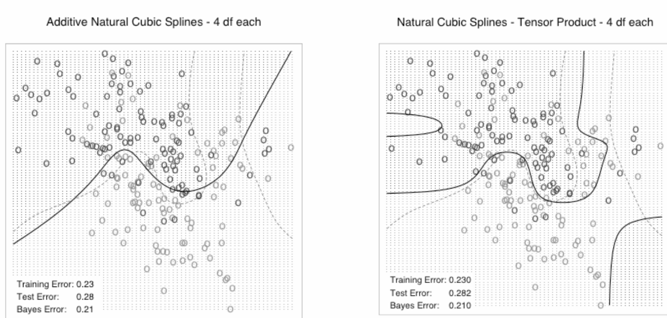
- Unlike 1D splines (which can be O(n)), complexity of thin plate splines is O(n^3) since there isn't sparse structure to exploit
- Can use fewer than N knots (lattice over domain)
- Using K knots reduces order to O(NK^2+K^3)

Additive Splines

- Function of the form
  $$f(X) = \alpha + f_1(X_1) + ... + f_d(X_d)$$
- Each fi is a univariate spline
- Assume f is additive and impose a penalty on each of the component functions

$$J[f] = J(f_1 + ... + f_d) = \sum_{j=1}^{d} f_j''(t_i)^2 dt_j$$

Additive v



ANOVA Spline Decomposition

- Extension of additive splines

$$f(X)=\alpha+\sum_j f_j(X_j)+\sum_{j<k} f_j(X_j,X_k)+\ldots$$

- Choices
  - Maximum order of interaction
  - Which terms to include
  - What representation
    - Regression splines with a small number of basis fns per coord
    - A complete basis as with smoothing splines

Questions

- A general question: can you give us a general method of how to choose the basis functions? What I mean does not refer to the model selection methods, like AIC, BIC, CV and etc, but refer to question such as have a quite skewed dataset, what kind of basis functions I should use to model this? (This might serve as a guidance to do AIC, BIC and etc instead of doing a full search in the mode space)
  - Yan Lui

Questions

- p. 140 Comparison of Additive Natural Cubic Splines to Natural Cubic Splines - Tensor Product. In text, "The tensor product basis can achieve more flexibility at the decision boundary, but introduces some spurious structure along the way."
- Based on the descriptions they give and the figures, I have trouble seeing when I'd desire the extra flexibility (and it isn't likely too hurt), or is it the case that the tensor product form is only useful in practice when included as part of a procedure like MARS?
  - Paul Bennet