# Basis expansions

Matthew S. Shotwell, Ph.D.

Department of Biostatistics
Vanderbilt University School of Medicine
Nashville, TN, USA

February 12, 2020

# Basics

- linear vs. nonlinear in $X$, $\theta$
- replace $X$ with transformations of $X$, then use linear model
- let $h_m(X) : \mathcal{R}^p \rightarrow \mathcal{R}$ for $m = 1 \ldots M$
- "linear basis expansion" in $X$
- then $f(X) = \sum_{m=1}^{M} \beta_m h_m(X)$
- given $h_m(X)$ the model is "linear" (i.e., in $\beta$)

# Examples

- $h_m(X) = X_m$ for $m = 1 \dots p$ - linear model
- $h_m(X) = I(L_m \leq X_k \leq U_m)$ - piecewise constant in $X_k$

# Controlling complexity

- let $\mathcal{D}$ be the "dictionary" of basis functions of size $|\mathcal{D}|$
- dictionary can be large; how to limit complexity?
- 1. decide beforehand, e.g., additivity (no interaction of inputs):

$$f(X) = \sum_{j=1}^{p} f_j(X_j) = \sum_{j=1}^{p} \sum_{m=1}^{M} \beta_{jm} h_{jm}(X_j)$$

- 2. selection methods, scan $\mathcal{D}$ for functions that improve fit, e.g., best subset, stepwise
- 3. regularization methods, e.g., lasso, ridge, elastic net, Bayesian methods

# Piecewise polynomials

- assume $X$ has support in $\mathcal{R}$
- divide $X$ into intervals defined by 'knots' $\xi_1, \xi_2, \ldots, \xi_K$
- $f(x)$ is polynomial in each interval

# Piecewise constant (0-degree polynomial)

Say there are two knots $\xi_1, \xi_2$ (three regions)

$$
\begin{aligned}
h_1(X) &= I(X < \xi_1) \\
h_2(X) &= I(\xi_1 \leq X < \xi_2) \\
h_3(X) &= I(\xi_2 \leq X)
\end{aligned}
$$

$$f(X) = \beta_1 I(X < \xi_1) + \beta_2 I(\xi_1 \leq X < \xi_2) + \beta_3 I(\xi_2 \leq X)$$

If $X$ is in first region ($X < \xi_1$) then $f(X) = \beta_1$. How many model parameters?

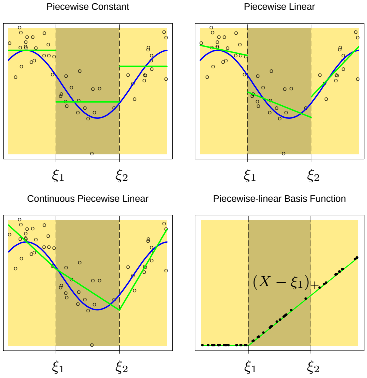



**FIGURE 5.1.** *The top left panel shows a piecewise constant function fit to some artificial data. The broken vertical lines indicate the positions of the two* knots *$\xi_1$ and $\xi_2$. The blue curve represents the true function, from which the data were generated with Gaussian noise. The remaining two panels show piecewise linear functions fit to the same data—the top right unrestricted, and the lower left restricted to be continuous at the knots. The lower right panel shows a piecewise–*

# Piecewise linear

$$
\begin{aligned}
h_1(X) &= I(X < \xi_1) \\
h_2(X) &= I(\xi_1 \leq X < \xi_2) \\
h_3(X) &= I(\xi_2 \leq X) \\
h_{3+m}(X) &= h_m(X)X \quad \text{for} \quad m = 1 \ldots 3
\end{aligned}
$$

How many model parameters?

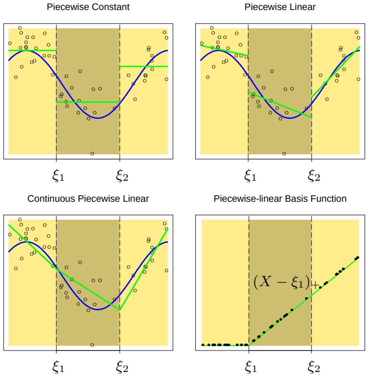



**FIGURE 5.1.** *The top left panel shows a piecewise constant function fit to some artificial data. The broken vertical lines indicate the positions of the two* knots *$\xi_1$ and $\xi_2$. The blue curve represents the true function, from which the data were generated with Gaussian noise. The remaining two panels show piecewise linear functions fit to the same data—the top right unrestricted, and the lower left restricted to be continuous at the knots. The lower right panel shows a piecewise–*

# Continuity restrictions

- can apply restrictions such that $f(X)$ is continuous at the region boundaries (i.e., the knots), e.g. in the piecewise linear model:

$$f(\xi_1^-) = f(\xi_1^+)$$
$$\beta_1 + \xi_1\beta_4 = \beta_2 + \xi_1\beta_5$$

$$f(\xi_2^-) = f(\xi_2^+)$$
$$\beta_2 + \xi_1\beta_5 = \beta_3 + \xi_2\beta_6$$

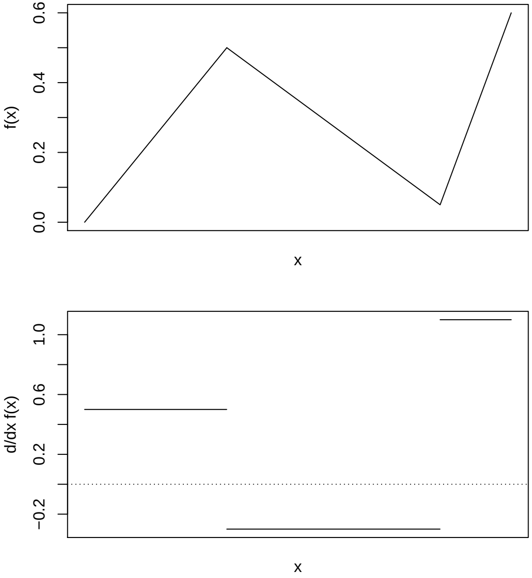
# Continuity in general

- if $f(x)$ can be traced without lifting pencil, then $f(x)$ is zero-order continuous
- if $f(x)$ and $f'(x)$ can be traced without lifting pencil, $f(x)$ is first-order continuous
- if $f(x)$ is piecewise linear, the constraint below enforces zero-order continuity

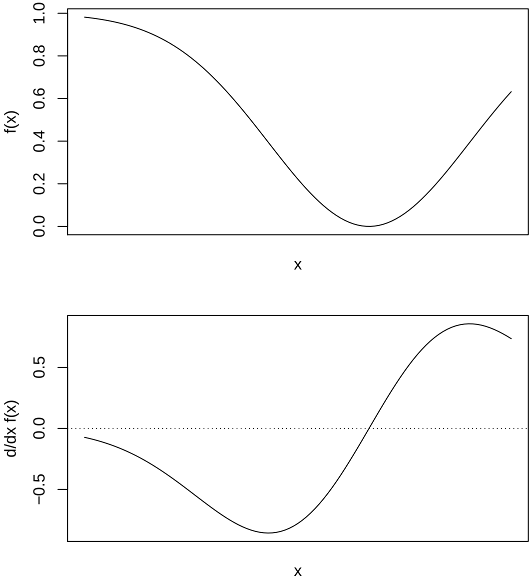
$$f(\xi_1^-) = f(\xi_1^+)$$
$$\beta_1 + \xi_1\beta_4 = \beta_2 + \xi_1\beta_5$$

$$f(\xi_2^-) = f(\xi_2^+)$$
$$\beta_2 + \xi_1\beta_5 = \beta_3 + \xi_2\beta_6$$

# zero-order continuous



# first-order continuous (at least)



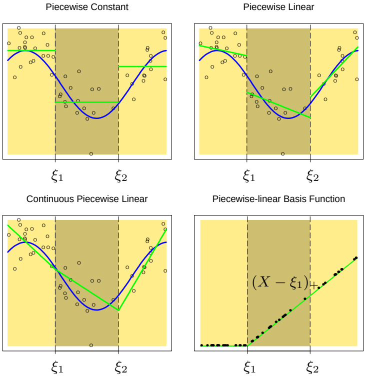



**FIGURE 5.1.** *The top left panel shows a piecewise constant function fit to some artificial data. The broken vertical lines indicate the positions of the two* knots *$\xi_1$ and $\xi_2$. The blue curve represents the true function, from which the data were generated with Gaussian noise. The remaining two panels show piecewise linear functions fit to the same data—the top right unrestricted, and the lower left restricted to be continuous at the knots. The lower right panel shows a piecewise–*

# Continuity restrictions

- two restrictions $\rightarrow$ two fewer parameters
- can write using $M = 4$ basis functions instead of $M = 6$

$$
\begin{aligned}
h_1(X) &= 1 \\
h_2(X) &= X \\
h_3(X) &= (X - \xi_1)_+ \\
h_4(X) &= (X - \xi_2)_+
\end{aligned}
$$

where $(a)_+ = a$ if $a > 0$ and $0$ otherwise

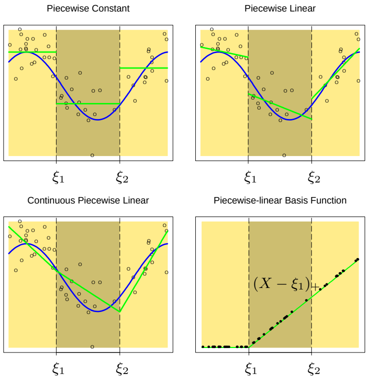



**FIGURE 5.1.** *The top left panel shows a piecewise constant function fit to some artificial data. The broken vertical lines indicate the positions of the two* knots *$\xi_1$ and $\xi_2$. The blue curve represents the true function, from which the data were generated with Gaussian noise. The remaining two panels show piecewise linear functions fit to the same data—the top right unrestricted, and the lower left restricted to be continuous at the knots. The lower right panel shows a piecewise–*

# Piecewise cubic polynomial

Where $\xi_0 = -\infty$ and $\xi_3 = \infty$, for $m = 1, 2, 3$

$$\begin{aligned} h_m(X) &= I(\xi_{m-1} \leq X < \xi_m) \\ h_{3+m}(X) &= h_m(X)X \\ h_{6+m}(X) &= h_{3+m}(X)X \\ h_{9+m}(X) &= h_{6+m}(X)X \end{aligned}$$

- if $K$ knots, then $(K+1) * 4$ function
- for $K = 2$, 12 functions, 12 parameters

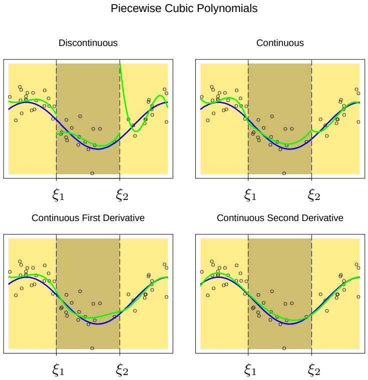



**FIGURE 5.2.** *A series of piecewise-cubic polynomials, with increasing orders of continuity.*

# Cubic spline

Piecewise cubic polynomial with continuity restrictions. For each knot $m = 1, 2$

$$
\begin{aligned}
f(\xi_m^-) &= f(\xi_m^+) \\
f'(\xi_m^-) &= f'(\xi_m^+) \\
f''(\xi_m^-) &= f''(\xi_m^+)
\end{aligned}
$$

- 3 restrictions per knot
- if $K$ knots, then $(K+1) * 4 - K * 3$ function
- for $K = 2$, 6 functions, 6 parameters

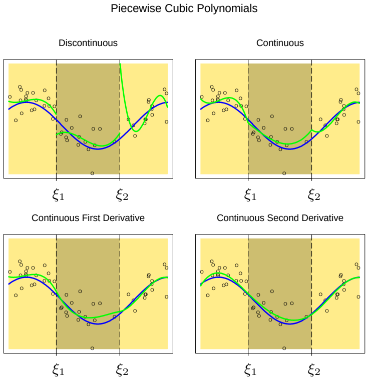



**FIGURE 5.2.** *A series of piecewise-cubic polynomials, with increasing orders of continuity.*

# Order-$M$ spline

- piecewise polynomial of order $M-1$
- continuous derivatives up to order $M-2$
- cubic spline is an order-4 spline
- most common are order-1 (piecewise constant), 2 (linear spline), and 4 (cubic spline)

# Natural cubic spline

- constrain to be linear beyond boundary knots
- four fewer parameters
- for knots $\xi_1, \ldots, \xi_K$

$$
\begin{aligned}
h_1(X) &= 1 \\
h_2(X) &= X \\
h_{2+k} &= d_k(X) - d_{K-1}(X) \\
d_k(X) &= \frac{(X-\xi_k)^3_+ - (X-\xi_K)^3_+}{\xi_K - \xi_k} \quad \text{for} \quad k = 1 \ldots K-2
\end{aligned}
$$

- if $K$ knots, then $(K+1) * 4 - K * 3 - 4 = K$ function
- $K$ basis functions for $K$ knots
- What happens when $K \leq 2$?