Statistical learning - Non-linear regression

Rita Almeida

PhD course 3045, VT 2018

Karolinska Institutet

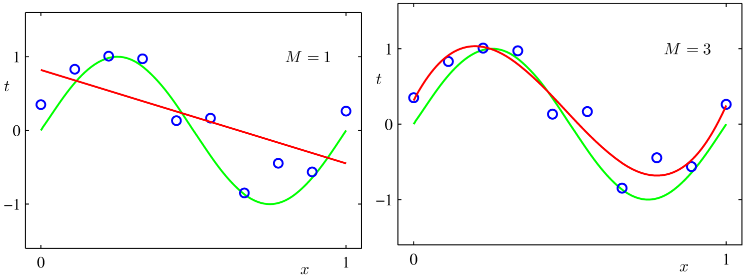
Polynomial regression

Linear regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Polynomial regression:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \cdots + \beta_p x_i^p + \epsilon_i$$



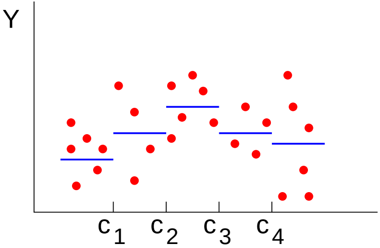
Step functions

Idea: break X into bins and fit a constant to each bin.

- X is transformed into a ordered categorical variable.

New variables:

$$
\begin{aligned}
C_0(X) &= I(X < c_1) \\
C_1(X) &= I(c_1 \leq X < c_2) \\
&\vdots \\
C_k(X) &= I(c_k \leq X)
\end{aligned}
$$



I the indicator function: 1 if the condition is true, 0 otherwise.

$$y_i = \beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \cdots + \beta_k C_k(x_i) + \epsilon_i$$

Basis functions

Regression can be done using other families of functions.

$$y_i = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \cdots + \beta_k b_k(x_i) + \epsilon_i$$

Special cases

- polynomial function: $b_j(x_i) = x_i^j$
- step function: $b_j(x_i) = I(c_j \leq x_i < c_{j+1})$

Regression splines

Idea: fit polynomials to different pieces of X, requiring the fits to be smooth.

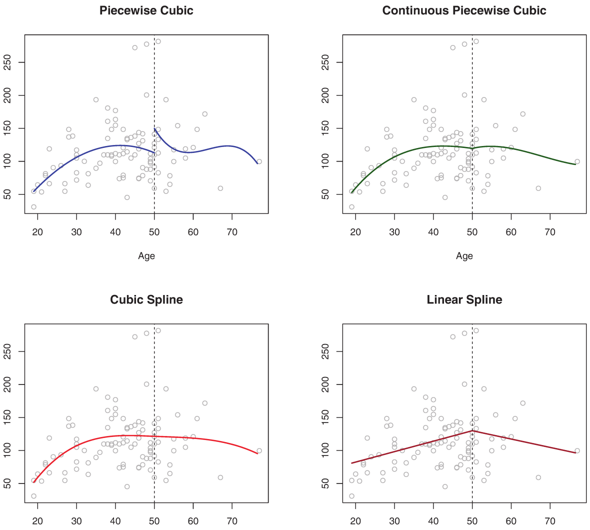


Figure adapted from James et al.

Regression splines

- Knots: points defining the pieces of X.
- For a polynomial with order p, constrain the polynomial and first $p-1$ derivatives to be continuous for all the knots.
- One can use a family of functions to represent a spline.

$$y_i = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \cdots + \beta_{k+p} b_{k+p}(x_i) + \epsilon_i$$

- For a polynomial of order p and for k knots one needs to use $p+k$ predictors (plus intercept), use $p+k+1$ degrees of freedom.

Example: cubic polynomial with 2 knots.

It can be implemented with the predictors: X, X^2, X^3, $h(X, \xi_1)$, $h(X, \xi_2)$, where ξ_1 and ξ_2 are the knots.

$$h(x, \xi) = \begin{cases} (x-\xi)^3 & \text{if } x > \xi \\ 0 & \text{otherwise} \end{cases}$$

Regression splines

- Natural splines: required to be linear at the outer pieces.
- Result in more stable estimates.

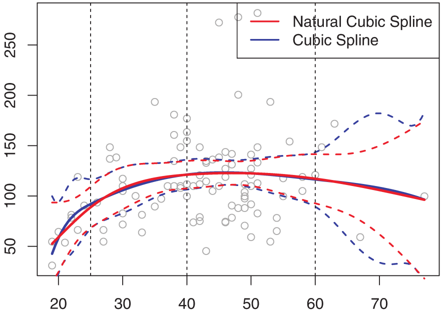


Figure adapted from James et al.

- There are different criteria to choose the knots.
- Number of knots can be determined by cross-validation.

Smoothing Splines

Idea: fitting the data by minimizing a function of the form *Loss + Penalty*, as for ridge regression and Lasso.

$$\sum_{i=1}(y_i - g(x_i))^2 + \lambda \int g''(t)^2 dt$$

- The function g that minimizes the expression is a smoothing spline.
- $g''(t)$ is a measure of roughness.
- When $\lambda \to \infty$, g is perfectly smooth, simple least squares fit.
- When $\lambda = 0$, g is very jumpy passing through all the points.
- λ controls the bias-variance trade-off.
- λ can be set by cross-validation.
- g will be a natural cubic spline, but not the same as obtained from the previous method.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 17 May 2017
doi: 10.3389/fnins.2017.00281

Check for updates

OPEN ACCESS

Edited by:
Raymond Scott Turner,
Georgetown University, United States

Reviewed by:

Time to Amyloid Positivity and Preclinical Changes in Brain Metabolism, Atrophy, and Cognition: Evidence for Emerging Amyloid Pathology in Alzheimer's Disease

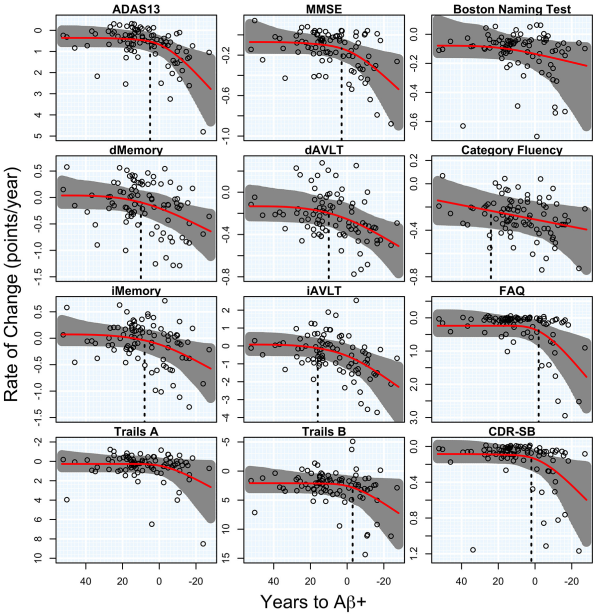
Philip S. Insel[1,2,3], Rik Ossenkoppele[4], Devon Gessert[5], William Jagust[6,7], Susan Landau[6,7], Oskar Hansson[1,8], Michael W. Weiner[2,3], and Niklas Mattsson[1,8,9] for the Alzheimer's Disease Neuroimaging Initiative[†]*

- 89 cognitivly normal subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database
- Individuals followed longitudinally up tp 10 years.
- β-amyloid (βA) estimates for at least 3 time points (for CSF or Florbetapir PET).
- The predictor was the estimated time to βA+, estimated for each individual using mixed models.
- Outcome variables: changes in cognitive tests / regional FDG-PET / regional volume (from MRI), estimated for each individual using mixed models, adjusting for age and gender, plus education/ . / scanner type and intracranial volume.

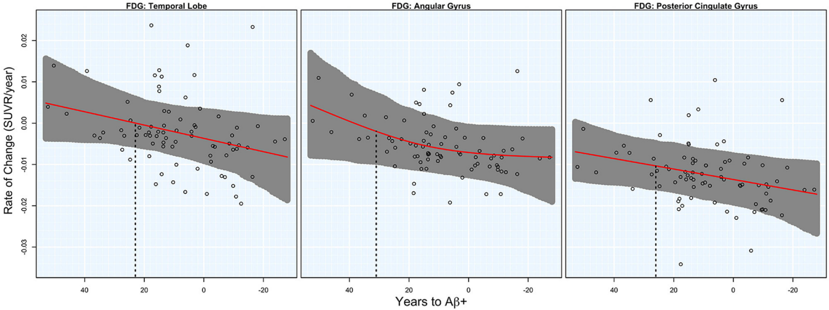
Cognitive tests:

Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale-cognitive subscale, 13-item version (ADAS13), immediate and delayed memory recall from the Wechsler Memory Scale (iMemory, dMemory), immediate and delayed Rey Auditory Verbal Learning Test (iAVLT, dAVLT), Trail Making Test parts A and B, Category Fluency, Boston Naming Test, Clinical Dementia Rating Sum of Boxes (CDR-SB), and the Functional Assessment Questionnaire (FAQ)

Example - results cognitive tests



Example - results FDG PET



Example - results volume MRI