**Generalized additive models (GAM)**

A GAM is a generalized linear model with a linear predictor involving a sum of smooth of functions of covariates. GAM in R uses the backfitting algorithm to combine different smoothing or fitting methods. The local scoring algorithm is used to iteratively fit weighted additive models by backfitting and Gauss-Seidel methods are used to fit additive models by iteratively smoothing partial residuals. The smoothing or fitting methods currently supported are local regression and smoothing splines.

**Advantages:**

The model allows for rather flexible specification of the dependence of the response on the covariates, but by specifying the model only in terms of ‘smooth functions’, rather than detailed parametric relationships.

**Notation:**

* X: a series of features, which specifies a linear predictor for response.
* Y: the response variable.
* : The smooth function.

**GAM in R:**

In general the model has a structure like

**A Gauss-Seidel method**: fitting additive models by iteratively smoothing partial residuals

Suppose the function is

1. Set for j=1,…,m.
2. Repeat steps 3 to 5 until the estimates, , stop changing.
3. For j=1,…,m repeat steps 4 and 5.
4. Calculate partial residuals:
5. Set equal to the result of smoothing with respect to .

**The smoothing or fitting methods:**  
**local regression**:

Combine much of the simplicity of linear least squares regression with the flexibility of nonlinear regression. It does this by fitting simple models to localized subsets of the data to build up a function that describes the deterministic part of the variation in the data, point by point.

Local regression model could also be known as locally weighted polynomial regression. The polynomial is fitted using weighted least squares, giving more weight to points near the point whose response is being estimated and less weight to points further away.

The traditional weight function is the tri-cube weight function

w(x)=

**Smoothing splines:**

The estimators perform a regularized regression over the natural spline basis, placing knots at all points Smoothing splines also circumvent the problem of knot selection, and simultaneously, they control for over-fitting by shrinking the coefficients of the estimated function.

Suppose the function is where are the truncated power basis functions for natural splines with knots at .(the natural splines are only defined for odd orders k)

The coefficients are chosen to minimize

,

where is the basis matrix defined as

,

and is the penalty matrix defined as

, i,j=1,…,n.

Given the optimal coefficients minimizing , the smoothing spline estimate at x is defined as

* The parameter is a tuning parameter, often called the smoothing parameter, and the higher the value of , the more shrinkage.
* Similar to least square regression, the is equal to
* Hence,