Cubic Splines

1 Method definition

A spline curve is built by using a different polynomial curve between each two data points. In other words, it is piecewise curve, made of pieces of different curves glued together. With cubic splines, the pieces are so well matched where they are glued that the gluing is not obvious.
A cubic smoothing spline behaves approximately like a kernel smoother, but it corresponds to the function $\hat{f}$ that minimizes the penalized residual sum of squares given by:

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

λ is the smoothing parameter, corresponding to the span in loess. The degree of freedom controls the amount of smoothing and corresponds to the trace of the smoothing matrix.

The generalized cross-validation is performed using this function:

$$CV(\lambda) = \frac{1}{n}\sum_{i=1}^{n}(y_i^* - \hat{f}_\lambda^{-i}(x_i))$$

Here $\hat{f}_\lambda^{-i}(x_i)$ is the leave-one-out smooth at x_i, that is constructed using all the data except for (x_i, y_i) and then the resulting least squares line is evaluated at x_i. CV is calculated for different values of λ and the λ that minimizes this criterion is chosen.

2 R code

The `calcRecomRates_cubicspline` (see below) allows to build the cubic spline model from the physical (`physcoord`) and genetical (`gencoord`) coordinates and to calculate recombination rates at some physical positions (`physpos`).

This function is based on the `R smooth.spline` function.
In `MareyMapOnline`, the generalized cross-validation is used, then the smoothing parameter and the degree of freedom are automatically estimated.
For more information about this method, write `?smooth.spline` in a `R` console.

```
calcRecomRates_cubicspline <- function(physpos, physcoord, gencoord) {
  modelCubicSpline <- smooth.spline(x = physCoord, y = genCoord)
  pp1 <- physpos + 1
  pm1 <- physpos - 1
  gp1 <- predict(modelCubicSpline, pp1)$y
  gm1 <- predict(modelCubicSpline, pm1)$y
  out <- mapply(function(xa, xb, ya, yb)
    {round((yb - ya) / ((xb - xa) / 1000000), 2)}, pm1, pp1, gm1, gp1)
  return(out)
}
```

3 Bibliography