The Registration Problem

Most analyzes only account for variation in amplitude.

Frequently, observed data exhibit features that vary in time.

Berkeley Growth Data

Aligning the curves reduces variation by 25%

Time-Warping Functions

Goal 1: transform observed curves $x_i(t)$ so that qualitative features line up

Requires a transformation of time.

Seek $s_i = w_i(t)$

so that $\tilde{x}_i(t) = x_i(s_i)$

are well aligned.

$w_i(t)$ are time-warping (also called registration) functions.

Option 1: Landmark Registration

We can try to align specific points.

for(i in 1:20){
    plot(agefine,Dhgtmat[,i])
    indexi = identify(agefine,D2hgtmat[,i],n=nmarks)
    hgtmarks[i,] = agefine[indexi]
}

Landmark registration

For each curve $x_i(t)$ we choose points $t_{i1}, \ldots, t_{iK}$

We need a reference (usually one of the curves)

$t_{01}, \ldots, t_{0K}$

so these define constraints

$w_i(t_{ij}) = t_{0j}$

Now we define a smooth function to go between these.
Identifying Landmarks

Consider identify to pick landmarks visually.
More automatically, major landmarks of interest:
- where \( x_i(t) \) crosses some value
- location of peaks or valleys
- location of inflections
Almost all are points at which some derivative of \( x_i(t) \) crosses zero.
In practise, zero-crossings can be found automatically, but usually still require manual checking.

Constraints on Warping Functions

Let \( t \in [0, T] \), the \( w_i(t) \) should follow a number of constraints:
- Initial conditions
  \[ w_i(0) = 0, \quad w_i(T) = T \]
- Landmarks
  \[ w_i(t_{ij}) = t_{0j} \]
- Monotonicity: if \( t_1 < t_2 \),
  \[ w_i(t_1) < w_i(t_2) \]

Computing the Landmark Registration

We compute \( w_i^{-1}(t) \) such that \( w_i^{-1}[w_i(t)] = t \).
1. Take \( t_{01}, \ldots, t_{0K} \) to be observation times
2. Response for \( w_i \) is \( t_{i1}, \ldots, t_{ik} \)
3. Smooth = usually close to interpolating the points
4. Obtain registered functions by re-inverting:
   \[ \tilde{x}_i(t) = x_i(w_i(t)) \]
Monotone smoothing – satisfies monotonicity requirements.
Frequently standard techniques also retain monotonicity.
Computationally cheaper, but should be checked.

In R

```r
hgtmarkmean = apply(hgtmarks, 2, mean)
WfdPar = fdPar(hgtbasis, 2, 1e-4)
landmarkres = landmarkreg(D2hgtffd, hgtmarks, hgtmarkmean, WfdPar)
regfd warpfd
```
A Comparison

Registered

Unregistered

Warping Functions

Result

Warping function below diagonal pushes registered function \textit{later} in time.

Automatic Methods

Landmark registration requires
- clearly identifiable landmarks
- manual care in defining and finding landmarks

\textbf{Goal 2:} Find registration functions to explain as much variation between curves as possible.
Squared Error for Shift Registration

Shift registration = just offset in time:

\[ w_i(t) = t + \delta_i \]

we can try to find \( \delta_1, \ldots, \delta_N \) to minimize between-curve sum of squares:

\[
\text{BCSSE}(\delta) = \sum_{i,j} \int (x_i(t + \delta_i) - x_j(t + \delta_j))^2 dt
\]

also requires a constraint, for example

\[
\sum \delta_i = 0
\]

Nonlinear optimization problem, but fairly manageable.

Why Squared Error Doesn’t Work for Flexible Methods

Amplitude-only variation is not ignored.

Re-thinking Registration

Major issue: we do not want to account for effects that are due solely to amplitude variation.

Idea: if I have \( y(t) = Ax(t) \), then bivariate plot of \( y(t), x(t) \) should be a straight line.

\[ A = 1 \Rightarrow \text{slope of 45\%, but we don’t care about this.} \]

Can test for how linear the relationship is.

Assessing Linearity

If \( y(t) \approx Ax(t) \), sample on points \( t_1, \ldots, t_K \) to get bivariate data.

The principle components decomposition should have one very large and one very small eigenvalue.

In functional terms, we try to minimize the second eigenvalue of

\[
T(w) = \left( \begin{array}{cc}
\int \{x_0(t)\}^2 dt & \int x_0(t)x[h(t)] dt \\
\int x_0(t)x[h(t)] dt & \int \{x[h(t)]\}^2 dt
\end{array} \right)
\]

so we define the (possibly penalized) criterion

\[
\text{MINEIG}_\lambda(w) = \xi_2(T(w)) + \lambda \int [LW(t)]^2 dt
\]
Some Notes

- Second eigenvalue = more general than maximizing correlation; allows multivariate functions
- numerical optimization and integration problem.
- Represent $w_i(t)$ as monotone smooth:
  \[ w_i(t) = C_0 + C_1 \int_0^t e^{W_i(s)} ds \]
- Opposite implications of sign of $W_i(t)$.
- Requires definition of reference function (should be able to generalize).
- No good cross-validation method.

Interpreting Registration with Monotone Smoothing

Recall that for monotone smoothing we have
\[ w_i(t) = C_0 + C_1 \int_0^t e^{W_i(s)} ds \]

Notes:
- $t > w_i(t)$ = events in $x_i(t)$ are running early
- $W_i(t) > 0$ ⇒ slope of $w_i(t) > 1$
- "registered time" speeding up $\approx$ events running behind template curve

From $W(t)$ to $w(t)$

```r
register.fd
mD2hgtfffd = mean(D2hgtfffd)
zerofd = fd(matrix(0, hgtbasis$nbasis, 20), hgtbasis)
WfdPar = fdPar(zerofd, int2Lfd(2), 1e-4)
regfd = register.fd(y0fd=mD2hgtfffd, yfd=D2hgtfffd, WfdPar)
```

```
regfd warpfd
```
Some Practical Notes

1. Remove obvious simple sources of variation first:
   - shifts in amplitude
   - shifts in phase
2. Zero-crossings are good landmarks when they are clear
3. Registering derivatives can be good practice, especially for further zero-crossing landmarks.
4. Landmark registration can then be refined by continuous registration.
5. Much recent work on competing methodologies.

Frequently suggested that warping functions may be of interest themselves (timing of events more important than amplitude).

- few examples of using these in practice
- $w_i(t)$ ill-defined, difficult to estimate well.
Summary

- Registration – important tool for analyzing non-amplitude variation.
- Easiest: landmark registration, requires manual supervision.
- Continuous registration: numerically difficult alternative, uncertain.
- Lots of unknowns – usually a preprocessing step.
- Warning: interaction with derivatives

\[
Dx[h(t)] = D[h(t)]D[x][h(t)]
\]

etc