Data and Dynamics

Fitting dynamic models directly to data is difficult

- non-linear relationships
- often non-deterministic systems (ie, lack of fit)
- need initial conditions

But functional data allows us to improve this

\[ Dx(t) = \beta(t)x(t) + \alpha(t)u(t) + \epsilon(t) \]

is just a concurrent linear model.

Partition of Variation

Use of stochastic systems now describes differences between observed curves into three

1. different solutions to \( Dx(t) = \beta(t)x(t) \) – different initial conditions.
2. different \( \epsilon_i(t) \) – random shocks to the system
3. different observational errors \( y_{ij} = x_i(t_{ij}) + \epsilon_{ij} \)

Accounting for the first two is equivalent to the techniques already discussed.

Some Notes

\[ Dx(t) = \beta(t)x(t) + \alpha(t)u(t) + \epsilon(t) \]

Model is not identifiable unless either

- restrictions are placed on \( \beta(t) \)
- we observe multiple copies of \( x(t) \).

Error term implies *stochastic* variation.

These are random disturbances that the system responds to much like it does to \( u(t) \).

However, the disturbances are different for each replication.

Lip Data

Measured position of lower lip saying the word “Bob”.

20 repetitions.

Result of considerable pre-processing; both smoothing and PCA

- initial rapid opening
- sharp transition to nearly linear motion
- rapid closure.

Approximate second-order model – think of lip as acting like a spring.

\[ D^2x(t) = \beta_1(t)Dx(t) + \beta_0(t)x(t) + \epsilon(t) \]
Looking at Derivatives

Principle Differential Analysis

For any homogeneous differential equation

\[ D^m x(t) = -\beta_{m-1}(t)D^{m-1}x(t) - \ldots - \beta_0(t)x(t) \]

we have an associated linear differential operator

\[ Lx = (D^m + \beta_{m-1}(t)D^{m-1} + \ldots + \beta_0 I)x \]

We also have a set of \( N \) functional observations \( x_i(t) \).

Process of estimating \( L \) is described as principle differential analysis (PDA) because of it’s resemblance to principle components analysis.

A 3-Dimensional Phase-Plane Plot

Fitting Criterion for PDA

Want to find an operator that comes as close as possible to satisfying

\[ Lx(t) \equiv 0, \ i = 1, \ldots, N \]

Requires estimating \( m \) functional parameters \( \beta_j(t) \), \( j = 1, \ldots, m - 1 \).

Regard \( Lx_i \) as the residual of the fit and use the squared error criterion

\[ SSE(L) = \sum \int [Lx_i(t)]^2 \ dt \]

If an input forcing function \( u_i \) has also been observed we fit

\[ SSE(L) = \sum \int [Lx_i(t) - u_i(t)]^2 \ dt \]

This is just a concurrent linear model.
PDA via fRegress

\[ \text{D}l_{\text{ipfd}} = \operatorname{deriv.fd}(l_{\text{ipfd}}, 1) \]
\[ \text{D}2l_{\text{ipfd}} = \operatorname{deriv.fd}(l_{\text{ipfd}}, 2) \]

\[ \text{bwtlist} = \text{list}(\text{fdPar}(\text{lipbasis}, 2, 0), \text{fdPar}(\text{lipbasis}, 2, 0)) \]
\[ \text{pdamod} = \text{fRegress}(\text{D}2l_{\text{ipfd}}, \text{list}(\text{D}l_{\text{ipfd}}, l_{\text{ipfd}}), \text{bwtlist}) \]

Confidence intervals measured for error associated with acceleration.

Confidence intervals measured for error associated with acceleration.

pda.fd, Lfd and cross validation

pda.fd conducts regression automatically, produces an Lfd object that measures error.

\[ \text{lambdas} = c(0, \exp(-15:-8)) \]

\[ \text{CVmat} = \text{matrix}(0, \text{length(lambdas)}, 20) \]

\[ \text{for}(i \text{ in } 1: \text{length(lambdas)}){ \}
\[ \text{bwtlist} = \text{list}(\text{fdPar}(\text{lipbasis}, 2, \text{lambdas}[i]), \]
\[ \text{fdPar}(\text{lipbasis}, 2, \text{lambdas}[i])) \]
\[ \text{for}(j \text{ in } 1:20){ \}
\[ \text{tpdaList} = \text{pda.fd}(\text{list}(\text{lipfd}[-j]), \text{bwtlist}) \]
\[ \text{tLfd} = \text{Lfd}(2, \text{tpdaList$bwtlist}) \]
\[ \text{tu} = \text{deriv.fd}(\text{lipfd}[j], \text{tLfd}) \]
\[ \text{CVmat}[i,j] = \text{inprod}(\text{tu}, \text{tu}) \]
\[ } \]

The Discriminant

Recall that

\[ d(t) = \left( \frac{\beta_1(t)}{2} \right)^2 + \beta_0 \]

\[ \text{dfd} = 0.25 \ast \text{pdaList$bwtlist[[2]]}$fd^2 - \text{pdaList$bwtlist[[1]]}$fd \]

- initial impulse
- middle period of damped behavior (vowel)
- around periods of undamped behavior with period around 30-40 ms.
On a Bifurcation Diagram

Plot \((-\beta_1(t), -\beta_0(t))\) from `pda.fd` and add the discriminant boundary.

Examining Residuals

Consider `pdaList$resfdlist` or `deriv.fd(lipfd, Lfd(2, pdaList$bwtlist)`

No Obvious Diagnostic Flaws

Can also measure over-all \(R^2\):

\[
R^2_{PDA} = 1 - \frac{\sum \int [Lx_i(t)]^2 dt}{\sum \int [D^m x_i(t)]^2 dt}
\]

Amount of total variation in \(m\)th derivative explained by relationship to others.

In this case it’s 0.96.

Comparison to Exact Solutions

Discrepancy is modeled between \(Lx(t)\) and 0.

We can also ask how well we recreate the original curves assuming \(Lx(t) \equiv 0\).

This asks to what extent variation in initial conditions explains differences between functions.
Comparison to 1st-Order Model

\[ Dx(t) = \beta_1(t)x(t) \]

Over-all \( R^2 \): 0.74

Estimating Differential Equations for Improved Smooths

We consider the melanoma data again:

- there appear to be a number of different time-scales
- a broad-scale linear trend
- a cyclic trend with an approximately 11-year cycle
- some potential changes in the amplitude of the cycle

We would like to target our smoothing to take advantage of these.

Designing a Linear Differential Operator

We want an \( \text{Lfd} \) that annihilates a combination of linear trend and sinusoid.

- Linear trends are set to zero by the operator
  \[ Lx(t) = D^2x(t) \]
- Sinusoidal trends with period \( \alpha \) are set to zero by
  \[ Lx(t) = D^2x(t) + \alpha x(t) \]
- Taking higher derivatives keeps these zero. Suggests an operator of the form
  \[ Lx(t) = D^4x(t) + \beta_1 D^3x(t) + \beta_2 D^2x(t) \]

we'll be a little more flexible and use

Estimating an Lfd Operator

1. Start with a smooth, penalizing the highest derivative
   \[
   \text{mfdPar} = \text{fdPar}(\text{mbasis}, \text{int2Lfd}(4), \lambda) \\
   \text{mfd2} = \text{smooth.basis}(\text{year}, \text{incidence}, \text{mfdPar}) \text{$\dollar $df}
   \]

2. Now estimate parameters in the \( \text{Lfd} \) by PDA or regression. In this case, we perform a PDA of the second derivative.
   \[
   \text{D2mfd} = \text{deriv.fd}(\text{mfd}, 2) \\
   \text{cbasis} = \text{create.constant.basis}(\text{range(melanoma$year)}) \\
   \text{pdafdPar} = \text{fdPar}(\text{cbasis}, 0, 0) \\
   \text{bwtlist} = \text{list(pdafdPar, pdafdPar)} \\
   \text{mpda} = \text{pda.fd}(\text{D2mfd}, \text{bwtlist})
   \]
Now we can re-smooth, note we need to put the zero components of the \( Lfd \) back in.

\[
\text{zfdfPar} = \text{fdPar(fd}(0,\text{cbasis}),0,0) \\
\text{zerolist} = \text{list(zfdfPar,zfdfPar)} \\
\text{mLfd} = \text{Lfd}(4,\text{c(zerolist,mpda\$bwtlist)}) \\
\text{mfdPar2} = \text{fdPar(mbasis,mlfd,lambda)} \\
\text{mfd2} = \text{smooth.basis(year,incidence,mfdPar2)}\text{\$fd}
\]

---

### Results on Melanoma Data

We can also compare some statistics:

<table>
<thead>
<tr>
<th></th>
<th>before</th>
<th>after</th>
</tr>
</thead>
<tbody>
<tr>
<td>gcv</td>
<td>0.094</td>
<td>0.0899</td>
</tr>
<tr>
<td>df</td>
<td>12.55</td>
<td>10.59</td>
</tr>
<tr>
<td>SSE</td>
<td>1.54</td>
<td>1.69</td>
</tr>
</tbody>
</table>

---

### Iterating the Concurrent Linear Model

Why stop at one? Take new smooth and repeat.

```r
for(j in 1:niter){
  D2mfd3 = deriv.fd(mfd3,2) 
  mpda3 = pda.fd(D2mfd3,bwtlist) 
  mLfd3 = Lfd(4,\text{c(zerolist,mpda\$bwtlist)}) 
  CVs = \text{rep}(0,\text{length(lambdas)}) 
  for(i in 1:length(lambdas)){
    lambda = lambdas[i] 
    mfdPar3 = \text{fdPar(mbasis,mlfd3,lambda)} 
    CVs[i] = \text{smooth.basis(year,incidence,mfdPar3)}\text{\$gcv}
  } 
  i = \text{which.min(CVs)} 
  lambda = lambdas[i] 
  mfdPar3 = \text{fdPar(mbasis,mlfd3,lambda)} 
  mfd3 = \text{smooth.basis(year,incidence,mfdPar3)}\text{\$fd}
}
```

---

### Evolution of a Smooth

Substantial initial drop in both followed by little change.
Multidimensional PDA

When there are $k$ dimensions, we search for $k$ linear differential operators.

With two dimensions $x(t)$ and $y(t)$ and a second-order equation we look operators to satisfy

$L_x x(t):$

$$\beta_0(t)x(t) + \beta_1(t)Dx(t) + \beta_2(t)D^2x(t) = \alpha_0(t)y(t) + \alpha_1(t)Dy(t)$$

$L_y y(t):$

$$\beta_0(t)y(t) + \beta_1(t)Dy(t) + \beta_2(t)D^2y(t) = \alpha_0(t)x(t) + \alpha_1(t)Dx(t)$$

as closely as possible.

This (should) happen in pda.fd automatically, but in the meantime we can coerce it.

Using `pda.fd` with the Handwriting Data

First, get the $fd$ object and derivative

```r
fdafd = smooth.basis(fdatime,fdaarray,fdaPar)$fd
Dfdafd = deriv.fd(fdafd,1)
```

Now, for $x$ treat $y$ as a forcing function and vice versa

```r
ufdlist1 = list(fdafd[,1],Dfdafd[,2])
ufdlist2 = list(fdafd[,2],Dfdafd[,1])
```

A smaller basis for the coefficients speeds things up.

Examining Residuals

Compare to original second derivatives to examine scale.

The Coefficients

Not immediately interpretable; but note that coefficients on $x$ and $y$ are much larger than those on derivatives.
Examining Stability

Recall that the stability of the system depends on the eigenvalues of

\[
\begin{pmatrix}
D^2x(t) \\
D^2y(t) \\
Dx(t) \\
Dy(t)
\end{pmatrix}
= \begin{pmatrix}
-\beta x_1(t) & -\alpha x_1(t) & -\beta x_0(t) & -\alpha x_0(t) \\
\alpha y_1(t) & -\beta y_1(t) & \alpha y_0(t) & -\beta y_0(t) \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
x(t) \\
y(t)
\end{pmatrix}
\]

Negative signs because we are measuring the \(\beta(t)\) relative to the Lfd instead of the differential equation.

Now we can take the eigen-decomposition at each point.

Stability Analysis of Handwriting Data

Conclusion: largely constant, cyclic behavior.

Summary

- `pda.fd` and `fRegress` provide access to the concurrent linear model
- Like PCA, PDA looks for a linear operator to explain variation
- Instead of variation among curves, PDA looks to explain variation among derivatives
- Retains physical interpretation
- Use of PDA to improve smooths (more later)
- Stability analysis in more than one dimension must be computed manually.