Clustering Classical Music Performances

Daniel J. McDonald

Department of Statistics, Indiana University

dajmcdon@indiana.edu, mypage.iu.edu/∼dajmcdon

Why?

- Easy to describe musical characteristics you like: “up-tempo,” “strong beat,” “good lyrics,” “jazzy,” etc.
- Harder to describe characteristics of a performance that you like.
- In classical music, there are hundreds or thousands of recordings of the same piece.
- Why do we like some better than others?

What’s different?

1. Mistakes
2. Extraneous noise
3. Recording quality
4. Articulation/Legato/Bowing/Breathing
5. Dynamics
6. Rubato/Tempo

The first three are mostly uninteresting, but the rest are about interpretation.

We like performances with “better” interpretations.

Piano music

With piano, can focus on dynamics and tempo

We have quantitative data on everything from a specially equipped piano.

This piano records keystroke velocity, pedaling, timing, duration.

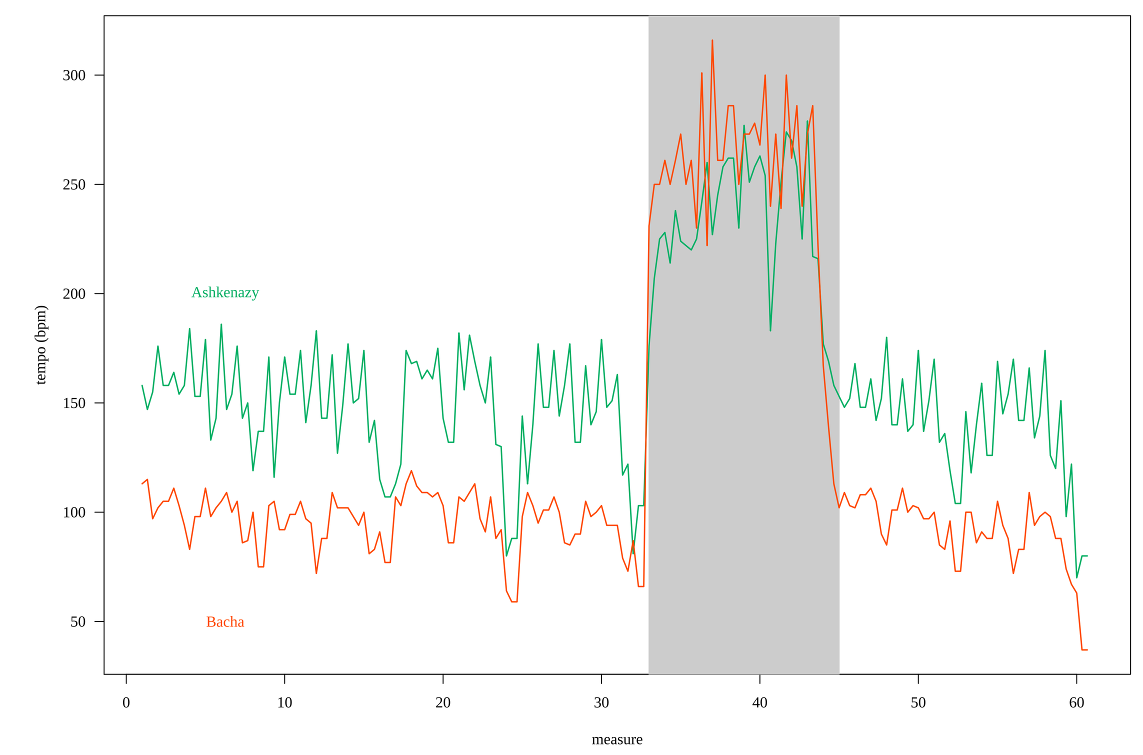
It lives in a studio at the IU Jacobs School of Music.

Data

Source and nature

- CHARM Mazurka Project
- Focus on timing only (dynamics also available)
- 50 recordings: Chopin Mazurka Op. 68 No. 3
- Recorded between 1931 and 2006
- 45 different performers

Functional data



Time points are highly correlated due to musical structure

Treat as functions $y_i(t)$ rather than vectors

Structure

measure	1	9	17	25	33	37	41	45	53
section	a	a′	b	a′	pedal	c	c′	a	a′

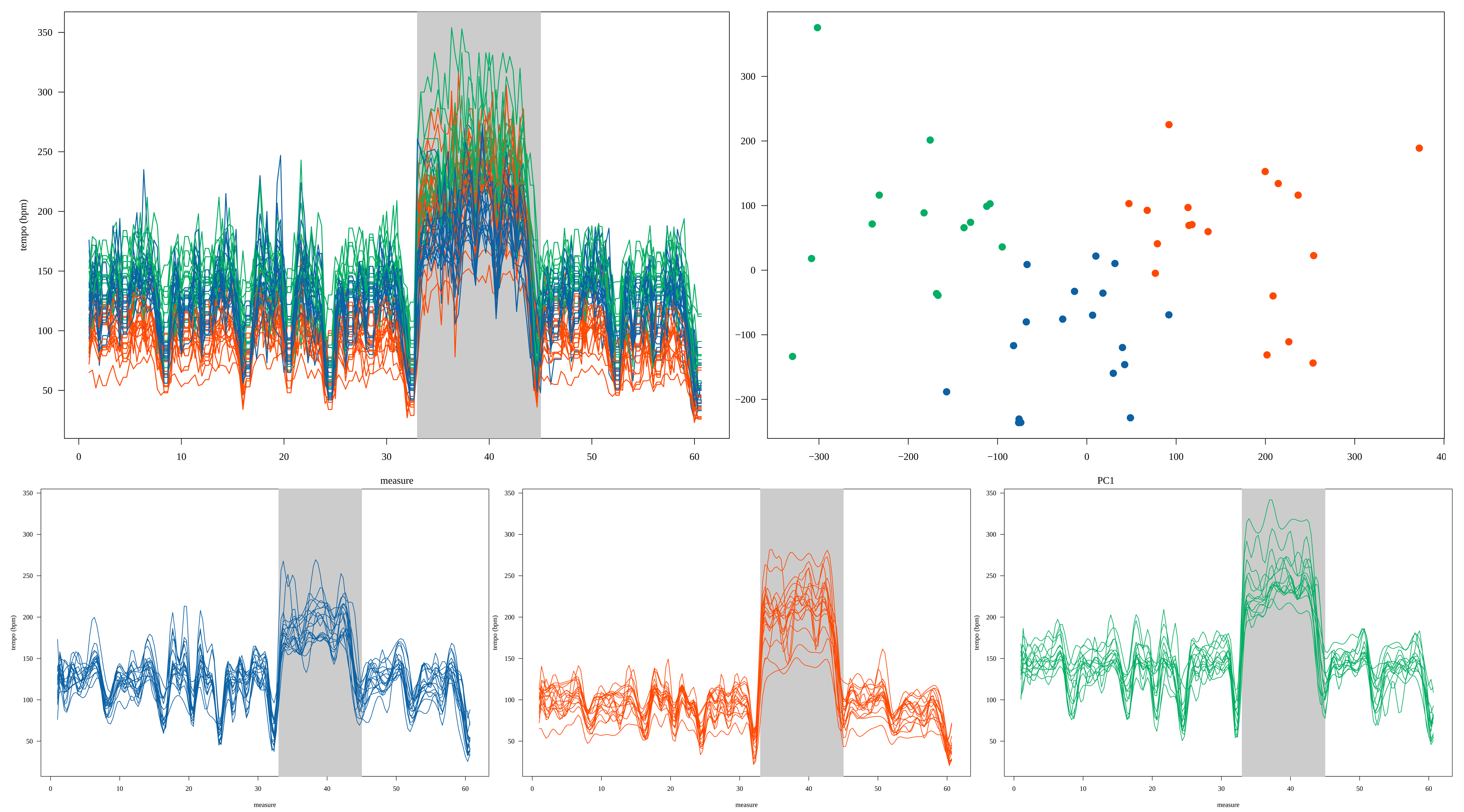
Preprocessing

- Regress functions on orthogonal basis
- Use estimated coefficients as feature vector
- “Decorolates” the temporal dependence

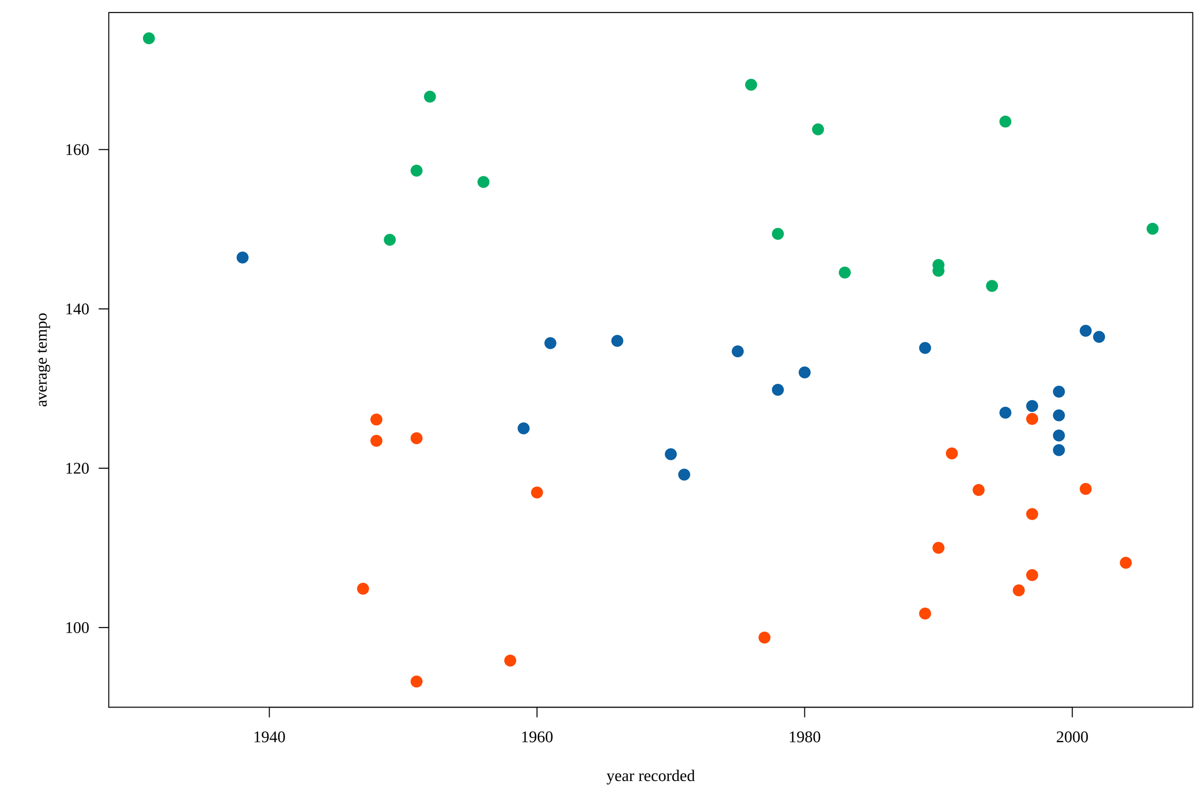
Methodological details

- Use Cubic B-splines evaluated at knots $t_1, \ldots, t_{\nu+4}$; take $x_i(t) = \sum_{j=1}^{\nu+4} \widehat{\beta}_{ij} B_j(t)$.
- Knot placement follows from musical structure
- Choose # of knots via GCV $\nu^* = \operatorname{argmin} \frac{1}{n} \sum_{k=1}^{n} \left(\frac{y_i(t_k) - x_i(t_k)}{1-\nu/n} \right)^2$; here $\nu^* = 60$: 1 knot/measure
- PCA on coefficient matrix $\Phi := (\widehat{\beta}_{ij}) = UDV^\top$ (functional PCA)
- Keep the leading eigenvectors and eigenvalues; used first 10 $U_{[1:10]}$
- K-means clustering on the projection $U_{[1:10]} D_{[1:10]}$, take $K = 3$

Recovered clusters



Chronological relationship?



Future work

- Using dynamic information
- Model based rather than nonparametric functional preprocessing
- Functional data on a manifold
  - Use Diffusion Map rather than PCA
  - Theoretical implications
- Musical structure recovery (Ren et. al. (JASA 2010))
  - Looking at tempo tracks for multiple recordings displays obvious structure
- Testing in real people