Statistical techniques in music psychology: An update

Daniel Müllensiefen
Goldsmiths, University of London

February 20, 2009

Factors that make music research difficult

1. Human diversity

Factors that make music research difficult

1. Human diversity
2. Model complexity

Factors that make music research difficult

1. Human diversity
2. Model complexity
3. Time dependency

Factors that make music research difficult

1. Human diversity
2. Model complexity
3. Time dependency
4. Prior knowledge of musical structure

Factors that make music research interesting

1. Human diversity - *Latent-class MDS and clustering*

Factors that make music research interesting

1. Human diversity - *Latent-class MDS and clustering*
2. Model complexity - *Classification and regression trees*

Factors that make music research interesting

1. Human diversity - *Latent-class MDS and clustering*
2. Model complexity - *Classification and regression trees*
3. Time dependency - *Functional Data Analysis*

Factors that make music research interesting

1. Human diversity - *Latent-class MDS and clustering*
2. Model complexity - *Classification and regression trees*
3. Time dependency - *Functional Data Analysis*
4. Prior knowledge of musical structure - *Bayesian and corpus-based models*

Empirical approaches for studying music

- Systematic Musicology: Aims at inducing laws and regularities in all music-related research areas (*The scientific study of music*

Empirical approaches for studying music

- Systematic Musicology: Aims at inducing laws and regularities in all music-related research areas (*The scientific study of music*
- Music Psychology / Music Cognition Research

Empirical approaches for studying music

- Systematic Musicology: Aims at inducing laws and regularities in all music-related research areas (*The scientific study of music*
- Music Psychology / Music Cognition Research
- Statistical Music Analysis

$\Longrightarrow$ All three approaches use statistics to formulate and test models

The standard repertoire - *not covered*

1. Descriptive statistics

The standard repertoire - *not covered*

1. Descriptive statistics
   Interential Statistics:

The standard repertoire - *not covered*

1. Descriptive statistics
   Interential Statistics:
2. Uni-variate comparisons between groups *(t-test, ANOVA)*

The standard repertoire - *not covered*

1. Descriptive statistics
   Interential Statistics:
2. Uni-variate comparisons between groups *(t-test, ANOVA)*
3. Bi-variate correlation (r) and association (χ^2) measures

The standard repertoire - *not covered*

1. Descriptive statistics
   Interential Statistics:
2. Uni-variate comparisons between groups *(t-test, ANOVA)*
3. Bi-variate correlation (r) and association (χ^2) measures
4. Multiple (linear) regression models

The standard repertoire - *not covered*

1. Descriptive statistics
   Interential Statistics:
2. Uni-variate comparisons between groups *(t-test, ANOVA)*
3. Bi-variate correlation (r) and association (χ^2) measures
4. Multiple (linear) regression models
5. Principal component analysis (PCA) and multi-dimensional scaling (MDS) for data reduction and visualisation

Related disciplines using advanced statistics - *not covered*

- Neuromusicology and neural models

Related disciplines using advanced statistics - *not covered*

- Neuromusicology and neural models
- Music Information Retrieval

Related disciplines using advanced statistics - *not covered*

- Neuromusicology and neural models
- Music Information Retrieval
- Individual differences and tests for musical ability

Multi-dimensional scaling (MDS)

In music psychology, MDS often used for judgemental dimensions in human similarity ratings for

- harmonic relationships (Krumhansl and Kessler 1982; Bharucha and Krumhansl 1983)

Multi-dimensional scaling (MDS)

In music psychology, MDS often used for judgemental dimensions in human similarity ratings for

- harmonic relationships (Krumhansl and Kessler 1982; Bharucha and Krumhansl 1983)
- timbre similarity (Kendall and Carterette 1991; Markuse and Schneider 1996)

Multi-dimensional scaling (MDS)

In music psychology, MDS often used for judgemental dimensions in human similarity ratings for

- harmonic relationships (Krumhansl and Kessler 1982; Bharucha and Krumhansl 1983)
- timbre similarity (Kendall and Carterette 1991; Markuse and Schneider 1996)
- stylistic judgements (Gromko 1993)

The classical MDS model

In the classical MDS model the distance between two objects j and j' on a small number of perceptual dimensions R is given by the the Euclidean distance $d_{jj'}$:

The classical MDS model

In the classical MDS model the distance between two objects j and j' on a small number of perceptual dimensions R is given by the the Euclidean distance $d_{jj'}$:

$$d_{jj'} = \left[\sum_{r=1}^{R} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

The classical MDS model

In the classical MDS model the distance between two objects j and j' on a small number of perceptual dimensions R is given by the the Euclidean distance $d_{jj'}$:

$$d_{jj'} = \left[\sum_{r=1}^{R} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

- Similarity judgements from different subjects are usually averaged

The INDSCAL model

The INDSCAL model (Carroll and Chang 1970) additionally models the different importance that N subjects might give to the R different dimensions by introducing $N \times R$ weight parameters w_{nr}:

$$d_{jj'} = \left[\sum_{r=1}^{R} w_{nr} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

The INDSCAL model

The INDSCAL model (Carroll and Chang 1970) additionally models the different importance that N subjects might give to the R different dimensions by introducing $N \times R$ weight parameters w_{nr}:

$$d_{jj'} = \left[\sum_{r=1}^{R} w_{nr} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

- But what do the individual weight parameters w_{nr} tell us?

The CLASCAL model

The CLASCAL model (Winsberg and De Soete 1993) groups the subjects into T latent classes and thus requires fewer weight parameters w_{tr}:

$$d_{jj't} = \left[\sum_{r=1}^{R} w_{tr} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

The CLASCAL model

The CLASCAL model (Winsberg and De Soete 1993) groups the subjects into T latent classes and thus requires fewer weight parameters w_{tr}:

$$d_{jj't} = \left[\sum_{r=1}^{R} w_{tr} (x_{jr} - x_{j'r})^2 \right]^{\frac{1}{2}}$$

- The latent classes might indicate different perceptual strategies with respect to a particular perceptual dimension

Applications

- McAdams et al. (1995) identified five different subject groups in a timbre similarity experiment

Applications

- McAdams et al. (1995) identified five different subject groups in a timbre similarity experiment
- Subject groups seem not related to musical training but to judgemental behaviour

Applications

- McAdams et al. (1995) identified five different subject groups in a timbre similarity experiment
- Subject groups seem not related to musical training but to judgemental behaviour
- Open question: What characterises different judgemental strategies?

Alternative ways of grouping similar subjects

Alternative ways of grouping similar subjects

1. Cluster subjects with hierarchical clustering methods (e.g. complete linkage, Ward) into few groups and then average group judgements (e.g. Pearce, Wiggins, and Müllensiefen, submitted)

Alternative ways of grouping similar subjects

1. Cluster subjects with hierarchical clustering methods (e.g. complete linkage, Ward) into few groups and then average group judgements (e.g. Pearce, Wiggins, and Müllensiefen, submitted)
2. Cluster subjects with probabilistic method (e.g. Gaussian Mixture) and then average their judgements

Typical issues that make data modelling complex

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables
- Many variables interact in a complex way with each other

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables
- Many variables interact in a complex way with each other
- Relationships between dependent and independent variable are be non-linear

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables
- Many variables interact in a complex way with each other
- Relationships between dependent and independent variable are be non-linear
- Distribution of data is *'anormal'*

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables
- Many variables interact in a complex way with each other
- Relationships between dependent and independent variable are be non-linear
- Distribution of data is *'anormal'*
- Visualising the statistical model is difficult

Typical issues that make data modelling complex

- Music perception task is influenced by many different variables
- Experimental or survey dataset has missing values for some observations on some variables
- Many variables interact in a complex way with each other
- Relationships between dependent and independent variable are be non-linear
- Distribution of data is *'anormal'*
- Visualising the statistical model is difficult

$\Longrightarrow$ Tree models might deal with these difficulties in a straightforward way

Classification and regression trees

- Tree models are well-known in machine learning but unknown in music cognition research

Classification and regression trees

- Tree models are well-known in machine learning but unknown in music cognition research
- Basic mechanism: Partition dataset recursively into smaller and more homogeneous subsets by predicting the dependent variable from the set of predictor variables

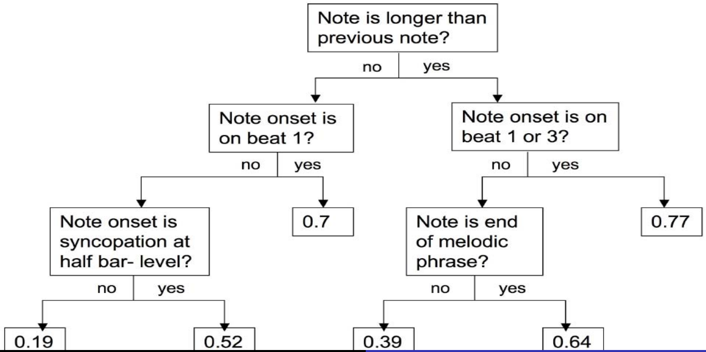
Classification and regression trees

- Tree models are well-known in machine learning but unknown in music cognition research
- Basic mechanism: Partition dataset recursively into smaller and more homogeneous subsets by predicting the dependent variable from the set of predictor variables
- Output is prediction of values on depedent variable for new cases

Classification and regression trees

- Tree models are well-known in machine learning but unknown in music cognition research
- Basic mechanism: Partition dataset recursively into smaller and more homogeneous subsets by predicting the dependent variable from the set of predictor variables
- Output is prediction of values on depedent variable for new cases
- Measures of model performance similar to linear / logistic models (R^2, classification error, x-validation error)

Example regression tree from Müllensiefen et al. (under review)



Applications

1.
    - Kopiez et al. (2006) predict performance in sight-reading task with tree (and other classification) models on basis of cognitive skills, practice skills, and practice hours

Applications

1.
   - Kopiez et al. (2006) predict performance in sight-reading task with tree (and other classification) models on basis of cognitive skills, practice skills, and practice hours
   - Results: Most effective predictors are trill speed, mental speed, time invested in sight-reading before 15

Applications

1.
   - Kopiez et al. (2006) predict performance in sight-reading task with tree (and other classification) models on basis of cognitive skills, practice skills, and practice hours
   - Results: Most effective predictors are trill speed, mental speed, time invested in sight-reading before 15
2.
   - Müllensiefen and Hennig (2006) explain music performance in memory task on basis of variables of musical structure and subjects' background
   - Results: Most important predictors are overall melodic similarity, similarity of accent structures, subjects' musical acitivty

Music perception unfolds in time

Where time is relevant:

- Performance parameters (e.g. tempo, timbre, loudness) change with time

Music perception unfolds in time

Where time is relevant:

- Performance parameters (e.g. tempo, timbre, loudness) change with time
- Reactions to music change over time

Music perception unfolds in time

Where time is relevant:

- Performance parameters (e.g. tempo, timbre, loudness) change with time
- Reactions to music change over time
- Simple averaging over time seems wrong

Music perception unfolds in time

Where time is relevant:

- Performance parameters (e.g. tempo, timbre, loudness) change with time
- Reactions to music change over time
- Simple averaging over time seems wrong
- Structure: Things close in musical time tend to correlate, things a distant times are more likely to be different

Objectives of Functional Data Analysis (FDA) (in music research)

- Analyse variables of interest as a function of time

Objectives of Functional Data Analysis (FDA) (in music research)

- Analyse variables of interest as a function of time
- Model discretely sampled data by smooth curve

Objectives of Functional Data Analysis (FDA) (in music research)

- Analyse variables of interest as a function of time
- Model discretely sampled data by smooth curve
- Look at *velocity* and *acceleration* of underlying processes

Objectives of Functional Data Analysis (FDA) (in music research)

- Analyse variables of interest as a function of time
- Model discretely sampled data by smooth curve
- Look at *velocity* and *acceleration* of underlying processes
- Break down process of interest into sub-processes that work on different time scales

Objectives of Functional Data Analysis (FDA) (in music research)

- Analyse variables of interest as a function of time
- Model discretely sampled data by smooth curve
- Look at *velocity* and *acceleration* of underlying processes
- Break down process of interest into sub-processes that work on different time scales
- Overcome disadvantes and inadequacies of conventional methods (repeated measurement and time series analysis)

Steps of a typical FDA

1. **Data gathering**: Sample continuos variable y at j discrete points in time

Steps of a typical FDA

1. **Data gathering**: Sample continuos variable y at j discrete points in time
2. **Smoothing and interpolation**: Model y_j by smooth latent process $x(t_j)$ and an error term ϵ_j.

$$y_j = x(t_j) + \epsilon_j$$

Steps of a typical FDA

1. **Data gathering**: Sample continuos variable y at j discrete points in time
2. **Smoothing and interpolation**: Model y_j by smooth latent process $x(t_j)$ and an error term ϵ_j.

$$y_j = x(t_j) + \epsilon_j$$

3. **Curve fitting**: Approximate smooth latent process with set of basis function ϕ_k (e.g. Fourier series, B-splines, wavelets)

$$x(t) = \sum_{k=1}^{K} c_k \phi_k(t)$$

Steps of a typical FDA *cont'd*

4. **Align smooth data of replications**: E.g. Align data from different subjects or several repetitions and possibly average

Steps of a typical FDA *cont'd*

4. **Align smooth data of replications**: E.g. Align data from different subjects or several repetitions and possibly average
5. **Display dynamic characteristics of process**: Differentiate data and plot 1st and 2nd derivative as *phase-plane* plots

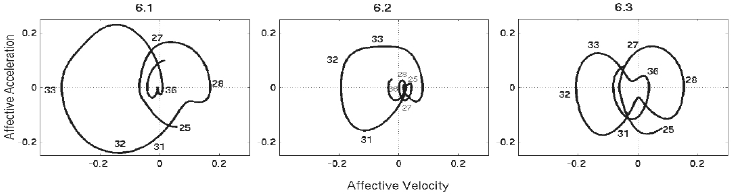


FIG. 6. (.1) Phase-plane plot for the auditory-only tension mean, showing Y' vs. Y'', or *affective velocity* vs. *affective acceleration*.

Steps of a typical FDA *cont'd*

6. **Model data**: Use functional linear modelling, principal component analysis (fPCA), ... to analyse smooth data, e.g.:

$$y_i(t) = \sum_{j=1}^{J} \beta_j(t) x_{ij} + \epsilon_i(t)$$

Applications of FDA in music research

1. Vines et al. (2005) and Vines et al. (2006) use FDA to study perceived tension during music listening

Applications of FDA in music research

1. Vines et al. (2005) and Vines et al. (2006) use FDA to study perceived tension during music listening
    - Result: Emotional *kinetic* and *potential* energy corresponds roughly compositional structure

Applications of FDA in music research

1. Vines et al. (2005) and Vines et al. (2006) use FDA to study perceived tension during music listening
   - Result: Emotional *kinetic* and *potential* energy corresponds roughly compositional structure
   - Tension perception differs for audio and/or audio-visual presentation of music

Applications of FDA in music research

1. Vines et al. (2005) and Vines et al. (2006) use FDA to study perceived tension during music listening
   - Result: Emotional *kinetic* and *potential* energy corresponds roughly compositional structure
   - Tension perception differs for audio and/or audio-visual presentation of music
2. Almansa and Delicado (ress) investigate the use of tempo as performative device across 28 recordings of Schumann's *Träumerei*

Applications of FDA in music research

1. Vines et al. (2005) and Vines et al. (2006) use FDA to study perceived tension during music listening
   - Result: Emotional *kinetic* and *potential* energy corresponds roughly compositional structure
   - Tension perception differs for audio and/or audio-visual presentation of music
2. Almansa and Delicado (ress) investigate the use of tempo as performative device across 28 recordings of Schumann's *Träumerei*
   - Result: Several meaningful components that are used differently by different performers (e.g. global tempo, ritardandi, contrast between phrases, change within phrases)

Two ways of applying Bayesian modelling

1. Inform model by prior knowledge available to the researcher

Two ways of applying Bayesian modelling

1. Inform model by prior knowledge available to the researcher
   A toy example:

Two ways of applying Bayesian modelling

1. Inform model by prior knowledge available to the researcher
   A toy example:
   - Goal: Model labeling of musical genres (A) from audio features (B) of a song (similar to Lippens et al. (2004))

Two ways of applying Bayesian modelling

1. Inform model by prior knowledge available to the researcher
   A toy example:
   - Goal: Model labeling of musical genres (A) from audio features (B) of a song (similar to Lippens et al. (2004))
   - Available information: *prior probability* $P(A)$, *conditional probability* $P(B|A)$ for a genre label A being correct if audio features B are present, and general probability of audio features $P(B)$

Two ways of applying Bayesian modelling

1. Inform model by prior knowledge available to the researcher
   A toy example:
   - Goal: Model labeling of musical genres (A) from audio features (B) of a song (similar to Lippens et al. (2004))
   - Available information: *prior probability* $P(A)$, *conditional probability* $P(B|A)$ for a genre label A being correct if audio features B are present, and general probability of audio features $P(B)$
   - Find genre with maximum likelihood given the audio features of a particular song according to Bayseian equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Two ways of applying Bayesian modelling

2. Inform model by prior knowledge available to subjects

Two ways of applying Bayesian modelling

2. Inform model by prior knowledge available to subjects
   Example: Sadakata et al. (2006)

Two ways of applying Bayesian modelling

2. Inform model by prior knowledge available to subjects
   Example: Sadakata et al. (2006)
   - Goal: Explain asymmetry in rhythm perception and production

Two ways of applying Bayesian modelling

2. Inform model by prior knowledge available to subjects
   Example: Sadakata et al. (2006)
   - Goal: Explain asymmetry in rhythm perception and production
   - In particular: How does the perceived rhythm class c relate to the time ratio t of two consecutive notes in a rhythm perception experiment?

$$P(c|t) = \frac{P(t|c)P(c)}{P(t)}$$

Two ways of applying Bayesian modelling

2. Inform model by prior knowledge available to subjects
   Example: Sadakata et al. (2006)
   - Goal: Explain asymmetry in rhythm perception and production
   - In particular: How does the perceived rhythm class c relate to the time ratio t of two consecutive notes in a rhythm perception experiment?

$$P(c|t) = \frac{P(t|c)P(c)}{P(t)}$$

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
    1. Theoretical (from Farey tree Peper et al. (1995))

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)
  3. Uniform

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)
  3. Uniform
  4. From production experiments

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)
  3. Uniform
  4. From production experiments
  5. Optimal priors from data fitting (only for comparion)

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)
  3. Uniform
  4. From production experiments
  5. Optimal priors from data fitting (only for comparion)
- Results: Empirical and theoretical priors give best results

Example: Sadakata et al. (2006) *cont'd*

- Use likelihood for production of duration ratio from existing experiments as $P(t|c)$
- How do different priors affect model accuracy? Priors used:
  1. Theoretical (from Farey tree Peper et al. (1995))
  2. Empirical (frequency counts from music corpora)
  3. Uniform
  4. From production experiments
  5. Optimal priors from data fitting (only for comparion)
- Results: Empirical and theoretical priors give best results
- $\Longrightarrow$ Even with vaguely musical stimuli listeners seem to use musical experience as frame of reference

The basic concept (Müllensiefen, Wiggins, and Lewis 2008)

- Observation about musical piece is only meaningful with reference to corpus of comparable pieces

The basic concept (Müllensiefen, Wiggins, and Lewis 2008)

- Observation about musical piece is only meaningful with reference to corpus of comparable pieces
- (Multivariate) distribution of features in music corpus is suitable level for aggregation of structures and comparison

The basic concept (Müllensiefen, Wiggins, and Lewis 2008)

- Observation about musical piece is only meaningful with reference to corpus of comparable pieces
- (Multivariate) distribution of features in music corpus is suitable level for aggregation of structures and comparison
- Features are cognitively relevant musical characteristics that can be extracted automatically

The basic concept (Müllensiefen, Wiggins, and Lewis 2008)

- Observation about musical piece is only meaningful with reference to corpus of comparable pieces
- (Multivariate) distribution of features in music corpus is suitable level for aggregation of structures and comparison
- Features are cognitively relevant musical characteristics that can be extracted automatically
- Frequent combinations of features describe general properties of musical style or culture

Example: Tversky similarity applied to melodies

Tversky's (1977) feature-based similarity for objects s and t:

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}$$

Example: Tversky similarity applied to melodies

Tversky's (1977) feature-based similarity for objects s and t:

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}$$

Adapted for measuring melodic similarity by

Example: Tversky similarity applied to melodies

Tversky's (1977) feature-based similarity for objects s and t:

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}$$

Adapted for measuring melodic similarity by

- defining interval chains (terms τ) as features of melodies

Example: Tversky similarity applied to melodies

Tversky's (1977) feature-based similarity for objects s and t:

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}$$

Adapted for measuring melodic similarity by

- defining interval chains (terms τ) as features of melodies
- defining salience function f as inverse of frequency of term τ in melody corpus C

$$IDF_C(\tau) = \log\left(\frac{|C|}{|m : \tau \in m|}\right)$$

Example: Tversky similarity applied to melodies

gives

Example: Tversky similarity applied to melodies

gives

$$\sigma(s,t) = \frac{\sum_{\tau \in s_n \cap t_n} IDF_C(\tau)}{\sum_{\tau \in s_n \cap t_n} IDF_C(\tau) + \alpha \sum_{\tau \in s_n \setminus t_n} IDF_C(\tau) + \beta \sum_{\tau \in t_n \setminus s_n} IDF_C(\tau)}$$

Example: Tversky similarity applied to melodies

Applications of the Tversky measure:

Example: Tversky similarity applied to melodies

Applications of the Tversky measure:

1. Evaluate similarity of melodies involved in plagiarism cases (Müllensiefen & Pendzich, in press)

Example: Tversky similarity applied to melodies

Applications of the Tversky measure:

1. Evaluate similarity of melodies involved in plagiarism cases (Müllensiefen & Pendzich, in press)
2. Measure increasing accuracy of mental representations of novel melodies (using recall paradigm by Sloboda and Parker (1985))

Software and references to start with

1. Latent-class MDS

   Software: Proprietary software (McAdams)

   Technical reference: Winsberg and De Soete (1993)

   Non-technical reference: McAdams et al. (1995)

2. Classification and Regression Trees

   Software: Libraries in R, Matlab, Statistica, Weka, SPSS (Clementine)

   Technical reference: Breiman et al. (1984)

   Non-technical reference: Witten and Frank (1999)

Software and references to start with

3. Functional Data Analysis

   Software: Libraries in R, Matlab
   Technical reference: Ramsay and Silverman (2005)
   Non-technical reference: Levitin et al. (2007)

4. Bayesian models

   Software: Libraries in R, Matlab, Weka, many specialised stand-alone applications
   Technical reference: MacKay (2003)
   Non-technical reference: Temperley (2007)

Almansa, J. and P. Delicado (2009, in press).
Analysing musical performance through functional data analysis: Rhythmic structure in schumann's träumerei.
Connection Science.

Bharucha, J. J. and C. L. Krumhansl (1983).
The representation of harmonic structure in music: Hierarchies of stability as a function of context.
Cognition 13, 63–102.

Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984).
Classification and Regression Trees.
Belmont (CA): Wadsworth.

Carroll, J. and J. Chang (1970, 283-319).
Analysis of individual difference sin multidimensional scaling via

an n-way generalization of eckart-young decomposition.
Psychometrika 52, 35.

Gromko, J. E. (1993).
Perceptual differences between expert and novice music listeners at multidimensional scaling analysis.
Psychology of Music 21, 34–47.

Kendall, R. A. and E. Carterette (1991).
Perceptual scaling of simultaneous wind instrument timbres.
Music Perception 8, 360–404.

Kopiez, R., C. Weihs, U. Ligges, and J. I. Lee (2006, October).
Classification of high and low achievers in a music sight-reading task.
Psychology of Music 34(1), 5–26.

Krumhansl, C. L. and E. Kessler (1982).
Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys.
Psychological Review 89, 334–368.

Levitin, D. J., R. L. Nuzzo, B. W. Vines, and J. O. Ramsay (2007).
Introduction to functional data analysis.
Canadian Psychology 48(3), 135–155.

Lippens, S., J.-P. Martens, T. De Mulder, and G. Tzanetakis (2004).
A comparison of human and automatic musical genre classification.
In *IEEE International Conference on Acoustics, Speech, and*

Signal Processing (ICASSP), Volume 4, pp. 223–236.

MacKay, D. (2003).
Information Theory, Inference, and Learning Algorithms.
Cambridge: Cambridge University Press.

Markuse, B. and A. Schneider (1996).
Ähnlichkeit, Nähe, Distanz: zur Anwendung multidimensionaler Skalierung in musikwissenschaftlichen Untersuchungen.
Systematische Musikwissenschaft/ Systematic Musicology/Musicologie systématique 4, 53–89.

McAdams, S., S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff (1995).
Perceptual sclaing of synthesized musical timbres: ommon dimensions, specificities, and latent subject classes.

Psychological Research 58, 177–192.

Müllensiefen, D. and C. Hennig (2006).
Modeling memory for melodies.
In M. Siliopoulou, R. Kruse, C. Borgelt, A. Nürnberger, and W. Gaul (Eds.), *From Data and Information Analysis to Knowledge Engineering*, Berlin, pp. 732–739. Gesellschaft für Klassifikation e.V.: Springer.

Müllensiefen, D., G. Wiggins, and D. Lewis (2008).
High-level feature descriptors and corpus-based musicology: Techniques for modelling music cognition.
In A. Schneider (Ed.), *Systematic and Comparative Musicology: Concepts, Methods, Findings*, Volume 24 of *Hamburger Jahrbuch für Musikwissenschaft*, pp. 133–155. Frankfurt: Peter Lang.

Peper, C., P. Beek, and P. van Wieringen (1995).
Multifrequency coordination in bimanual tapping: Asymmetrical coupling and signs of supercriticality.
Journal of Experimental Psychology: Human Perception and Performance 21, 117–1138.

Ramsay, J. O. and B. W. Silverman (2005).
Functional Data Analysis.
Berlin: Springer.

Sadakata, M., P. Desain, and H. Honing (2006).
The bayesian way to relate rhythm perception and production.
Music Perception 23(3), 269–288.

Sloboda, J. A. and D. H. Parker (1985).
Immediate recall of melodies.

In P. Howell, C. I., and R. West (Eds.), *Musical structure and cognition.*, pp. 143–167. London: Academic Press.

Temperley, D. (2007).
Music and Probability.
Cambridge, MA: MIT Press.

Vines, B. W., C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin (2006).
Cross-modal interactions in the perception of musical performance.
Cognition 101, 80–113.

Vines, B. W., R. L. Nuzzo, and D. J. Levitin (2005).
Analyzing temporal dynamics in music: Differential calculus, physics, and functional data analysis techniques.

Music Perception 23(2), 137–152.

Winsberg, S. and G. De Soete (1993).
A latent class approach to fitting the weighted euclidean model, CLASCAL.
Psychometrika 58, 315–330.

Witten, I. H. and E. Frank (1999).
Data mining: Practical machine learning tools and techniques with Java implementations.
San Francisco: Morgan Kaufmann.