A groupedData class

- designed to represent grouped data objects
- groupedData objects extend data.frames by incorporating a display formula that designates special roles for the variables and other attributes (e.g. labels, units)
- Most important special roles are: the response variable, the primary covariate, and the grouping factors.
- Other factors and covariates may also be assigned special roles
- methods available for displaying, summarizing, and modeling

Outer and Inner variables

- a variable is outer with respect to a grouping factor, if it is constant at each level of the factor (e.g. Sex in Orthodont)
- a variable is inner if it varies within the levels of the grouping factor (e.g. time in Theoph)
- a variable may be outer with respect to a factor and inner with respect to another (e.g. Length is inner to Site, but outer to Device in Wafer)
- inner variables are analogous to sub-plot factors and outer variables are analogous to whole-plot factors in split-plot designs.

Constructing groupedData objects

- A function to create an object of a given class is called a constructor for that class
- The groupedData constructor has two required arguments:
  - a display formula specifying the special roles for the variables of the form:

    `response ~ primary|groups`

    for single nesting, or

    `response ~ primary|g1/g2/../gN`

    for multiple nesting
  - a data.frame representing the data
- Optional arguments include: labels, units, outer, and inner

Plot methods

- Plot methods based on the Trellis library - any Trellis argument may be used
- classes and attributes are used to construct an adequate call to a Trellis function
- In the single-nesting case:
  - display and Trellis formulas coincide
  - By default, uses one panel per group - changed with outer argument
- In the multilevel case, may need to specify:
  - the display level
  - the collapsing level
  - the collapsing function

Grouped data - conclusions

- mixed-effects models are intended for grouped data
- data may be grouped according to single or multiple factors
- groupedData objects package information about roles of variables and other additional attributes with data
- plot methods operate as interface to Trellis displays, incorporating information about variables and attributes to generate a sensible plot
- gsummary summarizes groupedData objects and gapply applies function to subgroups of the data (defined by grouping factor(s))
- some modeling functions have methods for groupedData objects (e.g. lmList, lme)

Variance functions in lme

- varFunc classes are used to represent variance functions in lme
- constructor has the same name as the class
- customized classes may be added to the varFunc collection
- available varFunc classes
  - varFixed: fixed variance
  - varIdent: constant variance (per group)
  - varPower: power of covariate
  - varExp: exponential of covariates
  - varConstPower: constant plus power of covariate
  - varComb: combination of variance functions

Correlation structures

- model within-group correlation
- random effects account for some within-group correlation, but it may be necessary to also use correlation structures
- random effects structure and correlation structure may compete in model formulation: risk of overparameterization
- data collected over time tend to have serial correlation
- data collected over regions tend to present spatial correlation
- observations in different groups are assumed uncorrelated

Correlation Structures in lme

- corStruct classes are used to represent correlation structures in lme
- constructor has the same name as the class
- available corStruct classes
  - corAR1: AR(1)
  - corARMA: ARMA(p,q)
  - corBand: banded
  - corCAR1: continuous AR(1)
  - corCompSymm: compound symmetry
  - corExp: exponential
  - corGaus: Gaussian
  - corLin: linear
  - corSpher: spherical
  - corStrat: stratified
  - corSymm: general, unstructured

Classes of pd matrices in S

- pdMat classes are used to represent positive-definite matrices in lme
- constructor has the same name as the class and may be initialized in different ways
- customized classes may be added to the pdMat collection
- available pdMat classes
  - pdSymm: general pd matrix (default in lme)
  - pdDiag: diagonal
  - pdIdent: identity
  - pdCompSymm: compound symmetry
  - pdBlocked: block-diagonal
  - pdBand: banded
  - pdStrat: stratified