## A versatile tool for simulation of linear model data

Raju Rimal, Trygve Almøy and Solve Sæbø

Faculty of Chemistry and Bioinformatics (KBM)

Norwegian University of Life Scienes (NMBU)

12 Mar, 2018



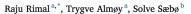
Contents lists available at ScienceDirect

### Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics



## A tool for simulating multi-response linear model data

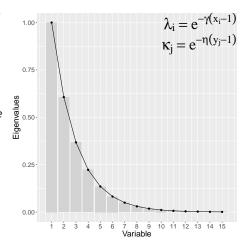




- a Faculty of Chemistry and Bioinformatics, Norwegian University of Life Sciences, As. Norway
- b Prorector, Norwegian University of Life Sciences, Ås, Norway

## Why Simrel

- Simulated data is used everywhere in research to compare methods, models, algorithms, techniques etc.
  Simrel can be a common tool for such purpose
- Simulate linear model data with wide range of properties using small set of tuning paramters, Example:
  - Controlling degree of multicollinearity in the simulated data
  - Specifying the relevant principle components for prediction

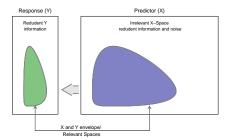


## Idea Behind

Reduction of regression Model: A Predictor sub-space (blue) is relevant for informative response sub-space (green)

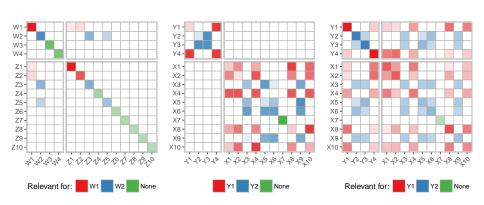
#### Relevant space within a model

A concept behind reduction of regression model

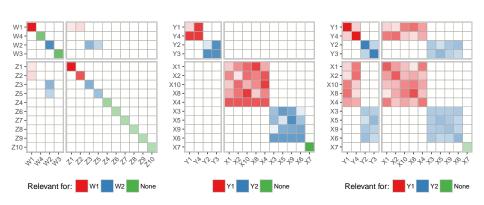


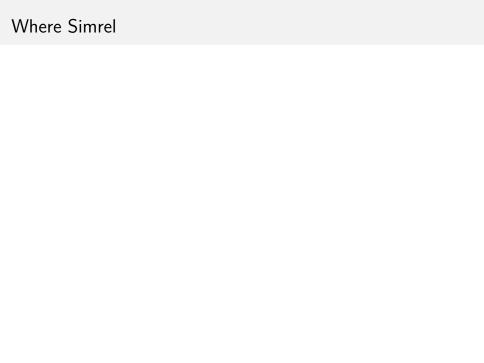
- A model defines its relationship with Response Space
- Subspace within these spaces (a reduced regression model) contains information for this relationship
- Set of orthogonal variables (Z) span the relevant predictor subspace (predictor components)
- Set of orthogonal variables (W) span the response subspace (response components)
- Implement this idea to construct the relevant covariance matrix and make simulation with it

## How it works



## How it works





# Example Simrel

# Acknoled gement

- Trygve
- Solve
- Franchisco
- Lars

