Modeling Nonlinear Moderation Effects in R

From Multiple Regression to Structural Equation Modeling

Tuo Liu Philipp Franikowski Andrea Hildebrandt

Exercise 1: The Impact of Employee Training Hours on Job Performance

Scenario of the `work` Dataset

You are investigating the effect of employee training hours (X) on job performance (Y). Preliminary observations suggest that the relationship between training hours and job performance may be nonlinear. Specifically, while some training is beneficial, excessive training might lead to diminishing returns or even negatively impact performance due to factors like cognitive overload or time away from core job responsibilities.

Variables

- X: Training hours per month (ranging from 0 to 20)
- Y: Job performance score (measured by performance reviews on a standardized scale)

Exercise 1.1: Nonlinear Regression Analysis

1. Fit a Regression Model:
    - Use a polynomial regression (e.g., a quadratic or a cubic polynomial regression) that is suited to model the potentially nonlinear relationship between training hours and job performance. Why did you choose **this** specification?
    - Do not forget about centering!
2. Model Specification (Polynomial):
    - Write down the equations of the fitted models with estimated coefficients.
3. Model Comparison / Assess Model Fit:

- Fit a simple linear regression model using the same data.
- Compare the polynomial models with the linear model using appropriate statistical procedures.

4. Variable (Predictor) Centering:

   a. Centering:
      - Do **not** center the variable X and M.
      - Estimate the correlation between the predictor and its high-order terms for both centered and uncentered variables.
      - Refit the polynomial regression models by using the uncentered predictor.

   b. Compare Results:

   - Compare the coefficients, standard errors, and model fit statistics between the centered and uncentered models.
   - Discuss any changes observed, i.e., explain how centering affects the interpretation of the intercept and the higher-order terms.

Exercise 1.2: Local Regression Analysis

1. Fit a Regression Model:

   - Use local regression (e.g., LOESS, span = 0.5) to model the potentially nonlinear relationship between training hours and job performance.
   - Do not forget about centering!

2. Model Comparison / Assess Model Fit:

   - Fit a linear and a quadratic regression model using the same data with `gam::gam()`.
   - Compare the local model to the polynomial models using appropriate statistical procedures.

Exercise 2: The Effect of Health Insurance Enrollment on Preventive Health Behavior, Moderated by Perceived Health Risk

Scenario of the `health` Dataset

You are studying how health insurance enrollment (X) influences preventive health behavior (Y), and how this effect is moderated by perceived health risk (M). Initial insights suggest that the relationship between health insurance and preventive behavior changes at different levels of perceived health risk and may involve nonlinear moderation effects.

Variables

- X: Health insurance enrollment (binary variable; sum-coded, i.e., -0.5 = No, 0.5 = Yes)
- M: Perceived health risk (continuous variable measured on a scale from 1 to 10)
- Y: Preventive health behavior score (higher values indicate greater engagement)

Exercise 2.1: Moderated Nonlinear Regression Analysis

1. Main Effect Analysis:
    a. Fit a Linear Regression Model:
        - Regress Y on X to test the linear effect of health insurance enrollment on preventive health behavior.
        - Report the estimated coefficient, its standard error, and p-value.
    b. Interpretation:
        - Interpret the effect of health insurance enrollment on preventive health behavior.
        - Discuss the significance of the findings.
2. Moderation Analysis:
    a. Testing for Moderation:
        - Examine whether perceived health risk (M) moderates the effect of X on Y.
        - Fit two models:
            – Model 1 (linear moderation) using linear regression
            – Model 2 (quadratic moderation) using quadratic regression
    b. Model Comparison:
        - Compare the two models using statistical tests.
        - Determine if nonlinear moderation model provides a significantly better fit than the linear moderation model.

3. Visualization:

    a. Plotting Moderation Effects:

    - Create a plot to visualize how the relationship (slope) between X and Y changes across different levels of M.

    b. Interpreting the Plots:

    - Describe the pattern observed in the plot.

4. Baseline Effect Analysis:

    a. Intercept as a Function of M:

    - Investigate whether the baseline level of preventive health behavior (when $X = 0$) is nonlinearly influenced by perceived health risk.

    b. Interpretation:

    - Explain how M affects Y independently of X.

Exercise 2.2: Moderated Nonlinear Nonparametric Regression Analysis

1. Moderation Analysis:

    a. Testing for Moderation:
    - Examine whether perceived health risk (M) moderates the effect of X on Y.
    - Fit the local moderation model using geographical weighted regression.

    b. Model Comparison:
    - Determine if local moderation of the slope provides a significantly better fit than quadratic moderation and linear moderation

2. Visualization:

    a. Plotting Moderation Effects:

    - Create a plot to visualize how the relationship (slope) between X and Y changes across different levels of M.

    b. Interpreting the Plot:

    - Describe the pattern observed in the plot.

3. Baseline Effect Analysis:

    a. Intercept as a Function of M:

    - Investigate with geographical weighted regression whether the baseline level of preventive health behavior (when $X = 0$) is nonlinearly influenced by perceived health risk.

b. Interpretation:

- Explain how M affects Y independently of X.

Exercise 3: The Effect of Conflict Framing on Media Trust, Moderated by Political Bias

Scenario of the `media` Dataset

You are investigating how Conflict Framing (X) influences Media Trust (Y), and whether this relationship is moderated by Political Bias (M). Both Conflict Framing and Media Trust are latent constructs measured by multiple items.

Theoretical Assumption:

- The influence of Conflict Framing on Media Trust is strongest among individuals with moderate political bias (i.e., those near the center of the political spectrum).
- For individuals with extreme political bias (either left or right), the influence is weaker because other framings are more important to them.

Variables

- Latent Variables:
  - ξ: Conflict Framing (measured by three items: `X1`, `X2`, `X3`)
  - η: Media Trust (measured by three items: `Y1`, `Y2`, `Y3`, `Y4`, `Y5`)
- Moderator:
  - M: Political Bias (measured on a continuous scale from -3 to 3, where negative values indicate left-leaning bias, positive values indicate right-leaning bias, and 0 indicates neutral)

Data Preparation and Measurement Invariance

- **Assumption:** Strong measurement invariance holds for the latent constructs across levels of M.

Exercise 3.1: Moderated Nonlinear SEM (MNSEM)

1. Model Specification and Fitting:
   a. Specify a Structural Equation Model (SEM):
      - Measurement Model: Define the measurement models for ξ and η using their respective items.
      - Structural Model: Specify that ξ predicts η

b. Scaling of Latent Variables:
   - Fit the SEM using different methods to scale the latent variables (e.g., UVI vs. ULI).
   - Compare the parameter estimates and model fit indices across different scaling methods.

2. Moderated Nonlinear SEM Analysis:

   a. Fit a Moderated Nonlinear SEM (MNSEM) Model:
      - Include the moderation effect of M on the relationship between ξ and η.
      - Test how the moderator M nonlinearly influences the latent regression coefficient between X and Y.
      - Write down the latent regression equation of the fitted model (MNSEM).
      - Explain why you chose this specific equation to capture the theorized inverted U-shaped relationship.
      - Interpret the estimated parameters from the final model, focusing on the moderation effect.
   b. Model Comparison:
      - Compare the moderated SEM model with a SEM model without moderation.
   c. Test for Quadratic Moderation:
      - Include both linear and quadratic term of M in the model to test if the moderation effect is quadratic.
      - Evaluate whether the inclusion of the quadratic term significantly improves the model.

3. Visualization:

   - Plot the Moderation Effect:
     – Create plots to visualize how the relationship between Conflict Framing and Media Trust changes across different levels of Political Bias, i.e., illustrate the nonlinear (quadratic) moderation effect.

Exercise 3.2: Local SEM (LSEM)

1. Local SEM Analysis:

   a. Fit a Local SEM (LSEM):
      - Include the moderation effect of M on the relationship between ξ and η.
      - Test how the moderator M nonlinearly influences the latent regression coefficient between X and Y.
      - Interpret the estimated parameters from the final model, focusing on the moderation effect.
   b. Model Comparison:

- Compare the moderated SEM model with a SEM model without moderation.

c. Test for Quadratic Moderation:

- Test for both a linear and quadratic moderation effect of M.
- Evaluate whether the inclusion of the quadratic term significantly improves the model.

2. Visualization:

- Plot the Moderation Effect:
  - Create plots to visualize how the relationship between Conflict Framing and Media Trust changes across different levels of Political Bias, i.e., illustrate the nonlinear moderation effect.